

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ :	A1	(11) International Publication Number: WO 96/37609 (43) International Publication Date: 28 November 1996 (28.11.96)
C12N 15/12, 15/85, 15/62, C07K 14/72, 19/00, C12N 5/10, A61K 38/16		
(21) International Application Number: PCT/GB96/01195		(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).
(22) International Filing Date: 20 May 1996 (20.05.96)		
(30) Priority Data:		
9510759.5 26 May 1995 (26.05.95) GB		
9513882.2 7 July 1995 (07.07.95) GB		
9517316.7 24 August 1995 (24.08.95) GB		
9605656.9 18 March 1996 (18.03.96) GB		
(71) Applicant (<i>for all designated States except US</i>): ZENECA LIMITED [GB/GB]; 15 Stanhope Gate, London W1Y 6LN (GB).		Published <i>With international search report.</i>
(72) Inventors; and		
(75) Inventors/Applicants (<i>for US only</i>): JEPSON, Ian [GB/GB]; 31 Gringer Hill, Maidenhead, Berkshire SL6 7LY (GB). MARTINEZ, Alberto [GB/GB]; Ivy Cottage, Terrace Road South, Binfield, Berkshire RG42 4DS (GB). GREENLAND, Andrew, James [GB/GB]; Tree Tops, Kingswood Court, Braywick Road, Maidenhead, Berkshire SL6 1DA (GB).		
(74) Agents: ROBERTS, Alison, Christine et al.; Zeneca Agrochemicals, Intellectual Property Dept., Jealott's Hill Research Station, P.O. Box 3538, Bracknell, Berkshire RG42 6YA (GB).		
(54) Title: A GENE SWITCH COMPRISING AN ECDYSONE RECEPTOR		
(57) Abstract		
		The invention relates to an insect steroid receptor protein which is capable of acting as a gene switch which is responsive to a chemical inducer enabling external control of the gene.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LJ	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

A gene switch comprising an ecdysone receptor

The present invention relates to the identification and characterisation of insect steroid receptors from the Lepidoptera species *Heliothis virescens*, and the nucleic acid encoding therefor. The present invention also relates to the use of such receptors, and such nucleic acid, particularly, but not exclusively, in screening methods, and gene switches.

By gene switch we mean a gene sequence which is responsive to an applied exogenous chemical inducer enabling external control of expression of the gene controlled by said gene sequence.

Lipophilic hormones such as steroids induce changes in gene expression to elicit profound effects on growth, cellular differentiation, and homeostasis. These hormones recognise intracellular receptors that share a common modular structure consisting of three main functional domains: a variable amino terminal region that contains a transactivation domain, a DNA binding domain, and a ligand binding domain on the carboxyl side of the molecule. The DNA binding domain contains nine invariant cysteines, eight of which are involved in zinc coordination to form a two-finger structure. In the nucleus the hormone-receptor complex binds to specific enhancer-like sequences called hormone response elements (HREs) to modulate transcription of target genes.

The field of insect steroid research has undergone a revolution in the last three years as a result of the cloning and preliminary characterisation of the first steroid receptor member genes. These developments suggest the time is ripe to try to use this knowledge to improve our tools in the constant fight against insect pests. Most of the research carried out on the molecular biology of the steroid receptor superfamily has been on *Drosophila melanogaster* (Diptera), see for example International Patent Publication No WO91/13167, with some in *Manduca* and *Galleria* (Lepidoptera).

It has been three decades since 20-hydroxyecdysone was first isolated and shown to be involved in the regulation of development of insects. Since then work has been carried out to try to understand the pathway by which this small hydrophobic molecule regulates a number of activities. By the early 1970s, through the studies of Clever and Ashburner, it was clear that at least in the salivary glands of third instar *Drosophila* larvae, the application of ecdysone lead to the reproducible activation of over a hundred genes. The ecdysone receptor in this pathway is involved in the regulation of two classes of genes: a small class (early genes) which are induced by the ecdysone receptor and a large class (late genes) which are repressed by the ecdysone receptor. The early class of genes are thought to have two functions reciprocal to those of the ecdysone receptor; the repression of the early transcripts and the induction of late gene transcription. Members of the early genes so far isolated and characterised belong to the class of molecules with characteristics similar to known

transcription factors. They are thus predicted to behave as expected by the model of ecdysone action (Ashburner, 1991). More recently, the early genes E74 and E75 have been shown to bind both types of ecdysone inducible genes (Thummel et al., 1990; Segraves and Hogness, 1991), thus supporting their proposed dual activities. It should be noted however, that the 5 activation of a hierarchy of genes is not limited to third instar larvae salivary glands, but that the response to the ecdysone peak at the end of larval life is observed in many other tissues, such as the imaginal disks (i.e. those tissues which metamorphose to adult structures) and other larval tissues which histolyse at the end of larval life (e.g. larval fat body). The model for ecdysone action as deduced by studying the third instar chromosome puffing may not apply 10 to the activation of ecdysone regulated genes in adults. In other words, the requirement for other factors in addition to the active ecdysone receptor must be satisfied for correct developmental expression (e.g. the *Drosophila* yolk protein gene expression in adults is under control of doublesex, the last gene in the sex determination gene hierarchy).

The ecdysone receptor and the early gene E75 belong to the steroid receptor 15 superfamily. Other *Drosophila* genes, including ultraspiracle, tailless, sevenup and FTZ-F1, also belong to this family. However, of all these genes only the ecdysone receptor is known to have a ligand, and thus the others are known as orphan receptors. Interestingly, despite the ultraspiracle protein ligand binding region sharing 49% identity with the vertebrate retinoic X receptor (RXR) ligand binding region (Oro et al., 1990), they do not share the 20 same ligand (i.e. the RXR ligand is 9-cis retinoic acid) (Heymann et al., 1992 and Mangelsdorf et al., 1992). All the *Drosophila* genes mentioned are involved in development, ultraspiracle for example, is required for embryonic and larval abdominal development. The protein products of these genes all fit the main features of the steroid receptor superfamily (Evans, 1988; Green and Chambon, 1988, Beato, 1989) i.e. they have a variable N terminus 25 region involved in ligand independent transactivation (Domains A and B), a highly conserved 66-68 amino acid region which is responsible for the binding of DNA at specific sites (Domain C), a hinge region thought to contain a nuclear translocation signal (Domain D), and a well conserved region containing the ligand binding region, transactivation sequences and the dimerisation phase (Domain E). The last region, domain F, is also very variable and 30 its function is unknown.

Steroid receptor action has been elucidated in considerable detail in vertebrate systems at both the cellular and molecular levels. In the absence of ligand, the receptor molecule resides in the cytoplasm where it is bound by Hsp90, Hsp70, and p59 to form the inactive complex (Evans, 1988). Upon binding of the ligand molecule by the receptor a conformational 35 change takes place which releases the Hsp90, Hsp70 and p59 molecules, while exposing the nuclear translocation signals in the receptor. The ligand dependent conformational change is seen in the ligand binding domain of both progesterone and retinoic acid receptors (Allan et

al., 1992a). This conformational change has been further characterised in the progesterone receptor and was found to be indispensable for gene transactivation (Allan et al., 1992b). Once inside the nucleus the receptor dimer binds to the receptor responsive element at a specific site on the DNA resulting in the activation or repression of a target gene. The receptor responsive elements usually consist of degenerate direct repeats, with a spacer between 1 and 5 nucleotides, which are bound by a receptor dimer through the DNA binding region (Domain C).

Whereas some steroid hormone receptors are active as homodimers others act as heterodimers. For example, in vertebrates, the retinoic acid receptor (RAR) forms heterodimers with the retinoic X receptor (RXR). RXR can also form heterodimers with the thyroid receptor, vitamin D receptor (Yu et al., 1991; Leid et al., 1992) and peroxisome activator receptor (Kliewer et al., 1992). Functionally the main difference between homodimers and heterodimers is increased specificity of binding to specific response elements. This indicates that different pathways can be linked, co-ordinated and modulated, and more importantly this observation begins to explain the molecular basis of the pleiotropic activity of retinoic acid in vertebrate development (Leid et al., 1992b). Similarly, the *Drosophila* ultraspiracle gene product was recently shown to be capable of forming heterodimers with retinoic acid, thyroid, vitamin D and peroxisome activator receptors and to stimulate the binding of these receptors to their target responsive elements (Yao et al., 1993). More significantly, the ultraspiracle gene product has also been shown to form heterodimers with the ecdysone receptor, resulting in cooperative binding to the ecdysone response element and capable of rendering mammalian cells ecdysone responsive (Yao et al., 1992). The latter is of importance since transactivation of the ecdysone gene alone in mammalian cells fails to elicit an ecdysone response (Koelle et al., 1991), therefore suggesting that the ultraspiracle gene product is an integral component of a functional ecdysone receptor (Yao et al., 1992). It is possible that the ultraspiracle product competes with other steroid receptors or factors to form heterodimers with the ecdysone receptor. Moreover it remains to be investigated if ultraspiracle is expressed in all tissues of the *Drosophila* larvae. Despite ultraspiracle being necessary to produce a functional ecdysone receptor, the mechanism by which this activation takes place is as yet undetermined.

We have now isolated and characterised the ecdysone steroid receptor from *Heliothis virescens* (hereinafter HEcR). We have found that surprisingly unlike the *Drosophila* ecdysone steroid receptor (hereinafter DEcR), in reports to-date, HEcR can be induced by known non-steroidal inducers. It will be appreciated that this provides many advantages for the system.

Steroids are difficult and expensive to make. In addition, the use of a non-steroid as the inducer allows the system to be used in agrochemical and pharmaceutical applications, not

least because it avoids application of a steroid which is already present in insects and/or mammals. For example, it would not be feasible to use a gene switch in a mammalian cell which was induced by a naturally occurring steroidal inducer. It will also be appreciated that for environmental reasons it is advantageous to avoid the use of steroids as inducers.

5 According to one aspect of the present invention there is provided DNA having the sequence shown in Seq ID No. 2, wherein Seq ID No 2 gives the sequence for the HEcR.

According to another aspect of the present invention there is provided DNA having part of the sequence shown in Seq ID No. 2, which encodes for the HEcR ligand binding domain.

10 According to another aspect of the present invention there is provided DNA having part of the sequence shown in Seq ID No. 2, which encodes for the HEcR DNA binding domain.

15 According to yet another aspect of the present invention there is provided DNA having part of the sequence shown in Seq ID No. 2, which encodes for the HEcR transactivation domain.

According to a further aspect of the present invention there is provided DNA having part of the sequence shown in Seq ID No. 2, which encodes for the HEcR hinge domain.

20 According to a still further aspect of the present invention there is provided DNA having part of the sequence shown in Seq ID No. 2, which encodes for the HEcR carboxy terminal region.

According to one aspect of the present invention there is provided DNA having the sequence shown in Seq ID No. 3, wherein Seq ID No 3 gives the sequence for the HEcR.

25 According to another aspect of the present invention there is provided DNA having part of the sequence shown in Seq ID No. 3, which encodes for the HEcR ligand binding domain.

According to another aspect of the present invention there is provided DNA having part of the sequence shown in Seq ID No. 3, which encodes for the HEcR DNA binding domain.

30 According to yet another aspect of the present invention there is provided DNA having part of the sequence shown in Seq ID No. 3, which encodes for the HEcR transactivation domain.

According to a further aspect of the present invention there is provided DNA having part of the sequence shown in Seq ID No. 3, which encodes for the HEcR hinge domain.

35 According to a still further aspect of the present invention there is provided DNA having part of the sequence shown in Seq ID No. 3, which encodes for the HEcR carboxy terminal region.

According to one aspect of the present invention there is provided DNA having the sequence shown in Seq ID No. 4, wherein Seq ID No 4 gives the sequence for the HEcR.

According to another aspect of the present invention there is provided DNA having part of the sequence shown in Seq ID No. 4, which encodes for the HEcR ligand binding domain.

According to another aspect of the present invention there is provided DNA having part of the sequence shown in Seq ID No. 4, which encodes for the HEcR DNA binding domain.

According to yet another aspect of the present invention there is provided DNA having part of the sequence shown in Seq ID No. 4, which encodes for the HEcR transactivation domain.

According to a further aspect of the present invention there is provided DNA having part of the sequence shown in Seq ID No. 4, which encodes for the HEcR hinge domain.

According to a still further aspect of the present invention there is provided DNA having part of the sequence shown in Seq ID No. 4, which encodes for the HEcR carboxy terminal region.

As mentioned above, steroid receptors are eukaryotic transcriptional regulatory factors which, in response to the binding of the steroid hormone, are believed to bind to specific DNA elements and activate transcription. The steroid receptor can be divided into six regions, designated A to F, using alignment techniques based on shared homology with other members of the steroid hormone receptor superfamily. Krust et al identified two main regions in the receptor, C and E. Region C is hydrophilic and is unusual in its high content in cysteine, lysine and arginine. It corresponds to a DNA-binding domain, sometimes referred to as the "zinc finger". It is the DNA binding domain which binds to the upstream DNA of the responsive gene. Such upstream DNA is known as the hormone response element or HRE for short. Region E is hydrophobic and is identified as the hormone (or ligand) binding domain. Region E can be further subdivided into regions E1, E2 and E3.

The region D, which separates domains C and E is highly hydrophobic and is flexible. It is believed that communication between domains E and C involves direct contact between them through region D, which provides a hinge between the two domains. Region D is therefore referred to as the hinge domain.

The mechanism of the receptor appears to require it to interact with some element(s) of the transcription machinery over and above its interactions with the hormone and the hormone response element. N-terminal regions A and B perform such a function and are jointly known as the transactivation domain. The carboxy-terminal region is designated F.

The domain boundaries of the HEcR can be defined as follows:

DOMAIN	INTERVALS	
	base pairs	amino acids
Transactivating (A/B)	114-600	1-162
DNA Binding (C)	601-798	163-228
Hinge (D)	799-1091	229-326
Ligand Binding (E)	1092-1757	327-545
C-Terminal End (F)	1758-1844	546-577

The DNA binding domain is very well defined and is 66 amino acids long, thus providing good boundaries. The above intervals have been defined using the multiple alignment for the ecdysone receptors (Figure 5).

The present invention also includes DNA which shows homology to the sequences of the present invention. Typically homology is shown when 60% or more of the nucleotides are common, more typically 65%, preferably 70%, more preferably 75%, even more preferably 80% or 85%, especially preferred are 90%, 95%, 98% or 99% or more homology.

The present invention also includes DNA which hybridises to the DNA of the present invention and which codes for at least part of the *Heliothis* ecdysone receptor transactivation domain, DNA binding domain, hinge domain, ligand binding domain and/or carboxy terminal region. Preferably such hybridisation occurs at, or between, low and high stringency conditions. In general terms, low stringency conditions can be defined as 3 x SCC at about ambient temperature to about 65°C, and high stringency conditions as 0.1 x SSC at about 65°C. SCC is the name of a buffer of 0.15M NaCl, 0.015M trisodium citrate. 3 x SSC is three time as strong as SSC and so on.

The present invention further includes DNA which is degenerate as a result of the genetic code to the DNA of the present invention and which codes for a polypeptide which is at least part of the *Heliothis* ecdysone receptor transactivation domain, DNA binding domain, hinge domain, ligand binding domain and/or carboxy terminal region.

The DNA of the present invention may be cDNA or DNA which is in an isolated form.

According to another aspect of the present invention there is provided a polypeptide comprising the *Heliothis* ecdysone receptor or a fragment thereof, wherein said polypeptide is substantially free from other proteins with which it is ordinarily associated, and which is coded for by any of the DNA of the present invention.

According to another aspect of the present invention there is provided a polypeptide which has the amino acid sequence of Seq ID No. 4 or any allelic variant or derivative thereof, wherein Seq ID No. 4 gives the amino acid sequence of the HEcR polypeptide.

According to another aspect of the present invention there is provided a polypeptide which has part of the amino acid sequence of Seq ID No. 4 or any allelic variant or derivative thereof, which sequence provides the HEcR ligand binding domain.

5 According to another aspect of the present invention there is provided a polypeptide which has part of the amino acid sequence of Seq ID No. 4 or any allelic variant or derivative thereof, which sequence provides the HEcR DNA binding domain.

According to yet another aspect of the present invention there is provided a polypeptide which has part of the amino acid sequence of Seq ID No. 4 or any allelic variant or derivative thereof, which sequence provides the HEcR transactivation domain.

10 According to a further aspect of the present invention there is provided a polypeptide which has the amino acid sequence of a part of Seq ID No. 4 or any allelic variant or derivative thereof, which sequence provides the HEcR hinge domain.

15 According to a still further aspect of the present invention there is provided a polypeptide which has the amino acid sequence of a part of Seq ID No. 4 or any allelic variant or derivative thereof, which sequence provides the HEcR carboxy terminal region.

For the avoidance of doubt, spliced variants of the amino acid sequences of the present invention are included in the present invention.

20 Preferably, said derivative is a homologous variant which has conservative amino acid changes. By conservation amino acid changes we mean replacing an amino acid from one of the amino acid groups, namely hydrophobic, polar, acidic or basic, with an amino acid from within the same group. An example of such a change is the replacement of valine by methionine and vice versa.

25 According to another aspect of the present invention there is provided a fusion polypeptide comprising at least one of the polypeptides of the present invention functionally linked to an appropriate non-*Heliothis* ecdysone receptor domain(s).

According to an especially preferred embodiment of the present invention the HEcR ligand binding domain of the present invention is fused to a DNA binding domain and a transactivation domain.

30 According to another embodiment of the present invention the DNA binding domain is fused to a ligand binding domain and a transactivation domain.

According to yet another embodiment of the present invention the transactivation domain is fused to a ligand binding domain and a DNA binding domain.

The present invention also provides recombinant DNA encoding for these fused polypeptides.

35 According to an especially preferred embodiment of the present invention there is provided recombinant nucleic acid comprising a DNA sequence encoding the HEcR ligand

binding domain functionally linked to DNA encoding the DNA binding domain and transactivation domain from a glucocorticoid receptor.

According to yet another aspect of the present invention there is provided recombinant nucleic acid comprising a DNA sequence comprising a reporter gene operably linked to a promoter sequence and a hormone response element which hormone response element is responsive to the DNA bonding domain encoded by the DNA of the present invention.

According to another aspect of the present invention there is provided a construct transformed with nucleic acid, recombinant DNA, a polypeptide or a fusion polypeptide of the present invention. Such constructs include plasmids and phages suitable for transforming a cell of interest. Such constructs will be well known to those skilled in the art.

According to another aspect of the present invention there is provided a cell transformed with nucleic acid, recombinant DNA, a polypeptide, or a fusion polypeptide of the present invention.

15 Preferably the cell is a plant, fungus or mammalian cell.

For the avoidance of doubt fungus includes yeast.

The present invention therefore provides a gene switch which is operably linked to a foreign gene or a series of foreign genes whereby expression of said foreign gene or said series of foreign genes may be controlled by application of an effective exogenous inducer.

20 Analogs of ecdysone, such as Muristerone A, are found in plants and disrupt the development of insects. It is therefore proposed that the receptor of the present invention can be used in plants transformed therewith as an insect control mechanism. The production of the insect-damaging product being controlled by an exogenous inducer. The insect-damaging product can be ecdysone or another suitable protein.

25 The first non-steroidal ecdysteroid agonists, dibenzoyl hydrazines, typified by RH-5849 [1,2-dibenzoyl, 1-tert-butyl hydrazide], which is commercially available as an insecticide from Rohm and Haas, were described back in 1988. Another commercially available compound in this series is RH-5992 [tebufenozide, 3,5-dimethylbenzoic acid 1-(1,1-dimethylethyl)-2(4-ethylbenzoyl) hydrazide]. These compounds mimic
30 20-hydroxyecdysone (20E) in both *Manduca sexta* and *Drosophila melanogaster*. These compounds have the advantage that they have the potential to control insects using ecdysteroid agonists which are non-steroidal. Further Examples of such dibenzoyl hydrazines are given in US Patent No. 5,117,057 to Rohm and Haas, and Oikawa et al, Pestic Sci, 41, 139-148 (1994). However, it will be appreciated that any inducer of the gene switch of the
35 present invention, whether steroid or non-steroidal, and which is currently or becomes available, may be used.

The gene switch of the present invention, then, when linked to an exogenous or foreign gene and introduced into a plant by transformation, provides a means for the external regulation of expression of that foreign gene. The method employed for transformation of the plant cells is not especially germane to this invention and any method suitable for the target plant may be employed. Transgenic plants are obtained by regeneration from the transformed cells. Numerous transformation procedures are known from the literature such as 5 agroinfection using *Agrobacterium tumefaciens* or its Ti plasmid, electroporation, microinjection or plants cells and protoplasts, microprojectile transformation, to mention but a few. Reference may be made to the literature for full details of the known methods.

Neither is the plant species into which the chemically inducible sequence is inserted 10 particularly germane to the invention. Dicotyledonous and monocotyledonous plants can be transformed. This invention may be applied to any plant for which transformation techniques are, or become, available. The present invention can therefore be used to control gene expression in a variety of genetically modified plants, including field crops such as canola, 15 sunflower, tobacco, sugarbeet, and cotton; cereals such as wheat, barley, rice, maize, and sorghum; fruit such as tomatoes, mangoes, peaches, apples, pears, strawberries, bananas and melons; and vegetables such as carrot, lettuce, cabbage and onion. The switch is also suitable for use in a variety of tissues, including roots, leaves, stems and reproductive tissues.

In a particularly preferred embodiment of the present invention, the gene switch of the 20 present invention is used to control expression of genes which confer resistance herbicide resistance and/or insect tolerance to plants.

Recent advances in plant biotechnology have resulted in the generation of transgenic 25 plants resistant to herbicide application, and transgenic plants resistant to insects. Herbicide tolerance has been achieved using a range of different transgenic strategies. One well documented example in the herbicide field is the use the bacterial xenobiotic detoxifying gene phosphinothricin acetyl transferase (PAT) from *Streptomyces hygroscopicus*. Mutated genes of plant origin, for example the altered target site gene encoding acetolactate synthase (ALS) 30 from *Arabidopsis*, have been successfully utilised to generate transgenic plants resistant to herbicide application. The PAT and ALS genes have been expressed under the control of strong constitutive promoter. In the field of insecticides, the most common example to-date is the use of the Bt gene.

We propose a system where genes conferring herbicide and/or insect tolerance would be expressed in an inducible manner dependent upon application of a specific activating chemical. This approach has a number of benefits for the farmer, including the following:

- 35 1. Inducible control of herbicide and/or insect tolerance would alleviate any risk of yield penalties associated with high levels of constitutive expression of herbicide and/or insect resistance genes. This may be a particular problem as early stages of growth

where high levels of transgene product may directly interfere with normal development. Alternatively high levels of expression of herbicide and/or insect resistance genes may cause a metabolic drain for plant resources.

2. The expression of herbicide resistance genes in an inducible manner allows the herbicide in question to be used to control volunteers if the activating chemical is omitted during treatment.
3. The use of an inducible promoter to drive herbicide and/or insect resistance genes will reduce the risk of resistance becoming a major problem. If resistance genes were passed onto weed species from related crops, control could still be achieved with the herbicide in the absence of inducing chemical. This would particularly be relevant if the tolerance gene confirmed resistance to a total vegetative control herbicide which would be used (with no inducing chemical) prior to sowing the crop and potentially after the crop has been harvested. For example, it can be envisaged that herbicide resistance cereals, such as wheat, might outcross into the weed wild oats, thus conferring herbicide resistance to this already troublesome weed. A further example is that the inducible expression of herbicide resistance in sugar beet will reduce the risk of wild sugar beet becoming a problem. Similarly, in the field of insect control, insect resistance may well become a problem if the tolerance gene is constitutively expressed. The use of an inducible promoter will allow a greater range of insect resistance control mechanisms to be employed.

This strategy of inducible expression of herbicide resistance can be achieved with a pre-spray of chemical activator or in the case of slow acting herbicides, for example N-phosphonomethyl-glycine (commonly known as glyphosate), the chemical inducer can be added as a tank mix simultaneously with the herbicide. Similar strategies can be employed for insect control.

This strategy can be adopted for any resistance conferring gene/corresponding herbicide combination, which is, or becomes, available. For example, the gene switch of the present invention can be used with:

1. Maize glutathione S-transferase (GST-27) gene (see our International Patent Publication No WO90/08826), which confers resistance to chloroacetanilide herbicides such as acetochlor, metolachlor and alachlor.
2. Phosphinotricin acetyl transferase (PAT), which confers resistance to the herbicide commonly known as glufosinate.
3. Acetolactate synthase gene mutants from maize (see our International Patent Publication No WO90/14000) and other genes, which confer resistance to sulphonyl urea and imadazolinones.

4. Genes which confer resistance to glyphosate. Such genes include the glyphosate oxidoreductase gene (GOX) (see International Patent Publication No. WO92/00377); genes which encode for 5-enolpyruvyl-3-phosphoshikimic acid synthase (EPSPS), including Class I and Class II EPSPS, genes which encode for mutant EPSPS, and genes which encode for EPSPS fusion peptides such as that comprised of a chloroplast transit peptide and EPSPS (see for example EP 218 571, EP 293 358, WO91/04323, WO92/04449 and WO92/06201); and genes which are involved in the expression of CPLyase.

5 Similarly, the strategy of inducible expression of insect resistance can be adopted for 10 any tolerance conferring gene which is, or becomes, available.

The gene switch of the present invention can also be used to controlled expression of foreign proteins in yeast and mammalian cells. Many heterologous proteins for many applications are produced by expression in genetically engineered bacteria, yeast cells and other eucaryotic cells such as mammalian cells.

15 As well as the obvious advantage in providing control over the expression of foreign genes in such cells, the switch of the present invention provides a further advantage in yeasts and mammalian cells where accumulation of large quantities of an heterologous protein can damage the cells, or where the heterologous protein is damaging such that expression for short periods of time is required in order to maintain the viability of the cells.

20 Such an inducible system also has applicability in gene therapy allowing the timing of expression of the therapeutic gene to be controlled. The present invention is therefore not only applicable to transformed mammalian cells but also to mammals *per se*.

25 A further advantage of the inducible system of the present invention in mammalian cells is that, because it is derived from a insect, there is less chance of it being effected by inducers which effect the natural mammalian steroid receptors.

30 In another aspect of the present invention the gene switch is used to switch on genes which produce potentially damaging or lethal proteins. Such a system can be employed in the treatment of cancer in which cells are transformed with genes which express proteins which are lethal to the cancer. The timing of the action of such proteins on the cancer cells can be controlled using the switch of the present invention.

The gene switch of the present invention can also be used to switch genes off as well as on. This is useful in disease models. In such a model the cell is allowed to grow before a specific gene(s) is switched off using the present invention. Such a model facilitates the study of the effect of a specific gene(s).

35 Again the method for producing such transgenic cells is not particularly germane to the present invention and any method suitable for the target cell may be used; such methods are known in the art, including cell specific transformation.

As previously mentioned, modulation of gene expression in the system appears in response to the binding of the HEcR to a specific control, or regulatory, DNA element. A schematic representation of the HEcR gene switch is shown in Figure 6. For ease of reference, the schematic representation only shows three main domains of the HEcR, namely 5 the transactivation domain, DNA binding domain and the ligand binding domain. Binding of a ligand to the ligand binding domain enables the DNA binding domain to bind to the HRE resulting in expression (or indeed repression) of a target gene.

The gene switch of the present invention can therefore be seen as having two components. The first component comprising the HEcR and a second component comprising 10 an appropriate HRE and the target gene. In practice, the switch may conveniently take the form of one or two sequences of DNA. At least part of the one sequence, or one sequence of the pair, encoding the HEcR protein. Alternatively, the nucleic acid encoding the HEcR can be replaced by the protein/ polypeptide itself.

Not only does the switch of the present invention have two components, but also one 15 or more of the domains of the receptor can be varied producing a chimeric gene switch. The switch of the present invention is very flexible and different combinations can be used in order to vary the result/to optimise the system. The only requirement in such chimeric systems is that the DNA binding domain should bind to the hormone response element in order to produce the desired effect.

20 The glucocorticoid steroid receptor is well characterised and has been found to work well in plants. A further advantage of this receptor is that it functions as a homodimer. This means that there is no need to express a second protein such as the ultraspiracle in order to produce a functional receptor. The problem with the glucocorticoid steroid receptor is that ligands used to activate it are not compatible with agronomic practice.

25 In a preferred aspect of the present invention the receptor comprises glucocorticoid receptor DNA binding and transactivation domains with a *Heliothis* ligand binding domain according to the present invention. The response unit preferably comprising the glucocorticoid hormone response element and the desired effect gene. In the Examples, for convenience, this effect gene took the form of a reporter gene. However, in non-test or non-30 screen situations the gene will be the gene which produces the desired effect, for example produces the desired protein. This protein may be a natural or exogenous protein. It will be appreciated that this chimeric switch combines the best features of the glucocorticoid system, whilst overcoming the disadvantage of only being inducible by a steroid.

In another preferred embodiment, the *Heliothis* ligand binding domain is changed, 35 and preferably replaced with a non-*Heliothis* ecdysone receptor ligand binding domain. For example, we have isolated suitable sequences from *Spodoptera exigua*.

Thus, according to another aspect of the present invention there is provided DNA having the sequence shown in Seq ID No. 6.

According to another aspect of the present invention there is provided DNA having part of the sequence shown in Seq ID No. 6, which encodes for the *Spodoptera* ecdysone ligand binding domain.

According to another aspect of the present invention there is provided DNA having part of the sequence shown in Seq ID No. 6, which encodes for the *Spodoptera* ecdysone hinge domain.

The present invention also provides the polypeptides coded for by the above DNA sequences of Seq ID No. 6.

A further advantage with such chimeric systems is that they allow you to choose the promoter which is used to drive the effector gene according to the desired end result. For example, placing the foreign gene under the control of a cell specific promoter can be particularly advantageous in circumstances where you wish to control not only the timing of expression, but also which cells expression occurs in. Such a double control can be particularly important in the areas of gene therapy and the use of cytotoxic proteins.

Changing the promoter also enables gene expression to be up- or down-regulated as desired.

Any convenient promoter can be used in the present invention, and many are known in the art.

Any convenient transactivation domain may also be used. The transactivation domain VP16 is a strong activator from Genentech Inc., and is commonly used when expressing glucocorticoid receptor in plants. Other transactivation domains derived for example from plants or yeast may be employed.

In a preferred embodiment of the present invention, the DNA binding domain is the glucocorticoid DNA binding domain. This domain is commonly a human glucocorticoid receptor DNA binding domain. However, the domain can be obtained from any other convenient source, for example, rats.

According to another aspect of the present invention there is provided a method of selecting compounds capable of being bound to an insect steroid receptor superfamily member comprising screening compounds for binding to a polypeptide or fusion polypeptide of the present invention, and selecting said compounds exhibiting said binding.

According to another aspect of the present invention there is provided a compound selected using the method of the present invention.

According to another aspect of the present invention there is provided an agricultural or pharmaceutical composition comprising the compound of the present invention.

According to yet another aspect of the present invention there is provided the use of the compound of the present invention as a pesticide, pharmaceutical and/or inducer of the switch. It will be appreciated that such inducers may well be useful as insecticides in themselves.

5 According to a further aspect of the present invention there is provided a method of producing a protein or peptide or polypeptide comprising introducing into a cell of the present invention, a compound which binds to the ligand binding domain in said cell.

Various preferred features and embodiments of the present invention will now be described by way of non-limiting example with reference to the accompanying examples and 10 figures, in which figures:

Figure 1 (Sequence ID No. 1) shows the DNA sequence amplified from first strand cDNA made from mRNA isolated from *Heliothis virescens* Fourth instar larvae. The underlined sequences refer to the position of the degenerate oligonucleotides. At the 5' end the sequence matches that of the oligonucleotide while at the 3' end 12 nucleotides of the 15 original oligonucleotide are observed;

Figure 2 (Sequence ID No. 2) shows the DNA sequence contained within the clone pSK19R isolated from a random primed cDNA *Heliothis virescens* library; Sequence is flanked by EcoRI sites;

Figure 3 (Sequence ID No. 3) shows the DNA sequence contained within the clone 20 pSK16.1 isolated from a random primed cDNA *Heliothis virescens* library ;

Figure 4 (Sequence ID No. 4) DNA sequence of 5'RACE products (in bold) fused to sequence of clone pSK16.1. The ORF (open reading frame) giving rise to the *Heliothis virescens* ecdysone receptor protein sequence is shown under the corresponding DNA sequence;

25 Figure 5 (Sequence ID No. 5) shows the protein sequence alignment of the ecdysone receptors DmEcR (*Drosophila melanogaster*), CtEcR (*Chironomus tentans*), BmEcR (*Bombyx mori*), MsEcR (*Manduca sexta*), AaEcR (*Aedes aegypti*) and HvEcR (*Heliothis virescens*). "*" indicates conserved amino acid residue. ":" indicates a conservative amino acid exchange;

30 Figure 6 shows a model of an embodiment of the glucocorticoid/*Heliothis* ecdysone chimeric receptor useable as a gene switch;

Figure 7 shows a plasmid map of the clone pcDNA319R. The three other mammalian expression vectors were constructed in the same way and look similar but for the size of the insert;

35 Figure 8 shows a plasmid map of the reporter construct used to analyse the activity of the *Heliothis virescens* ecdysone receptor;

Figure 9 is a graph which shows the effect of Muristerone A and RH5992 in reporter activity in HEK293 cells co-transfected with pcDNA3H3KHEcR alone (filled bars) or with α RXR (stripped bars);

5 Figure 10 shows a plasmid map of the Maize expression vector containing the Glucocorticoid receptor (HG1 or pMF6HG1PAT);

Figure 11 shows a plasmid map of the maize expression vector containing the chimeric glucocorticoid/*Drosophila* ecdysone receptor pMF6GREcRS;

Figure 12 shows a plasmid map of the maize expression vector containing the chimeric glucocorticoid/*Heliothis* ecdysone receptor pMF6GRHEcR;

10 Figure 13 shows a plasmid map of the plant reporter Plasmid containing the glucocorticoid response elements fused to the -60 S35CaMV promoter fused to GUS, p221.9GRE6;

15 Figure 14 shows a plasmid map of the plant reporter plasmid containing the glucocorticoid response elements fused to the -46 S35CaMV promoter fused to GUS, p221.10GRE6;

Figure 15 shows a graph showing the effect of Muristerone A and Dexamethasone in Maize AXB protoplasts transformed with pMF6HG1PAT (GR) and p221.9GRE6 (reporter);

20 Figure 16 shows a graph showing the effect of Muristerone A and Dexamethasone in Maize AXB protoplasts transformed with pMF6GREcRS (effector) and p221.9GRE6 (reporter);

Figure 17 shows a graph showing the effect of Muristerone A and Dexamethasone in Maize AXB protoplasts transformed with pMF6GRHEcR (effector) and p221.9GRE6 (reporter);

25 Figure 18 shows a graph showing the effect of RH5849 in Maize AXB protoplasts transformed with pMF6GREcRS (effector) and p221.9GRE6 (reporter);

Figure 19 shows a graph showing the effect of RH5992 in Maize AXB protoplasts transformed with pMF6GREcRS (effector) and p221.9GRE6 (reporter);

Figure 20 shows a graph showing the effect of RH5992 in Maize AXB protoplasts transformed with pMF6GRHEcR (effector) and p221.9GRE6 (reporter);

30 Figure 21 shows a graph which shows the dose response effect of RH5992 in Maize AXB protoplasts transformed with pMF6GRHEcR (effector) and p221.9GRE6 (reporter);

Figure 22 shows a plasmid map of the tobacco expression vector containing the chimeric glucocorticoid/ *Drosophila* ecdysone receptor, pMF7GREcRS;

35 Figure 23 shows a plasmid map of the tobacco expression vector containing the chimeric glucocorticoid/ *Heliothis* ecdysone receptor, pMF7GRHEcR;

Figure 24 shows a graph which shows the effect of RH5992 in Tobacco mesophyll protoplasts transformed with pMF6GRHEcR (Effector) and p221.9GRE6 (reporter);

- Figure 25 shows a plasmid map of the mammalian expression vector containing the chimeric glucocorticoid/*Heliothis* ecdysone receptor, pcDNA3GRHEcR;
- Figure 26 shows a plasmid map of the reporter plasmid pSWGRE4;
- Figure 27 shows a graph which shows a RH5992 dose response curve of CHO cells transfected with pcDNA3GRHEcR and pSWGRE4;
- 5 Figure 28 shows a graph which shows the effect of Muristerone A and RH5992 on HEK293 cells co-transfected with pcDNA3GRHEcR and pSWGRE4;
- Figure 29 shows a plasmid map of the binary vector ES1;
- Figure 30 shows a plasmid map of the binary vector ES2;
- 10 Figure 31 shows a plasmid map of the binary vector ES3;
- Figure 32 shows a plasmid map of the binary vector ES4;
- Figure 33 shows a plasmid map of the effector construct TEV-B112 made to express the HEcR ligand binding domain in yeast;
- Figure 34 shows a plasmid map of the effector construct TEV8 made to express the 15 HEcR ligand binding domain in yeast;
- Figure 35 shows a plasmid map of the effector construct TEVVP16-3 made to express the HEcR ligand binding domain in yeast;
- Figure 36 shows a plasmid map of the mammalian expression vector containing the chimeric glucocorticoid VP16/*Heliothis* ecdysone receptor, pcDNA3GRVP16HEcR;
- 20 Figure 37 shows a plasmid map of the maize expression vector containing the chimeric glucocorticoid VP16/*Heliothis* ecdysone receptor, pMF6GRVP16HEcR;
- Figure 38 shows a plasmid map of the maize expression vector containing the chimeric glucocorticoid VP16/*Heliothis* ecdysone receptor, pMF7GRVP16HEcR;
- Figure 39 shows a graph which shows the effect of RH5992 in Maize AXB 25 protoplasts transformed with pMF6GRVP16HEcR (effector) and p221.9GRE6 (reporter);
- Figure 40 (Sequence ID No. 6) shows the DNA sequence of the hinge and ligand binding domains of the *Spodoptera exigua* ecdysone receptor;
- Figure 41 (Sequence ID No. 7) shows the protein sequence alignment of the *Heliothis* 19R and *Spodoptera* SEcR *Taq* clone hinge and ligand binding domains. "*" indicates 30 conserved amino acid residue. ":" indicates a conservative amino acid exchange;
- Figure 42 shows a graph which shows the effect of RH5992 on Tobacco mesophyll protoplasts transformed with pMF7GRHEcR (effector) and either p221.9GRE6 (Horizontal strips) or p221.10GRE6 (vertical strips).

Example I - Cloning of the *Heliothis* Ecdysone Receptor**A. Probe generation**

- 5 The rational behind the generation of the probe to isolate *Heliothis* homologues to the steroid/thyroid receptor superfamily members was based on comparing the sequences of developmentally regulated steroid/thyroid receptor superfamily members. The sequences available showed a highly conserved motif within the DNA binding domain of the RAR and THR (thyroid) receptors. The motifs were used to design degenerate oligonucleotides for
- 10 PCR amplification of sequences derived from cDNA template produced from tissue expected to express developmentally regulated steroid/thyroid receptor superfamily members (ie. larval tissues).

The sense oligonucleotide is based on the peptide sequence CEGCKGFF which at the DNA level yields an oligonucleotide with degeneracy of 32 as shown below :

15 ZnFA5' 5' TGC GAG GGI TGC AAG GAI TTC TT 3'
 T A T A T

The antisense oligonucleotide is based on the reverse complement nucleotide sequence derived from the peptide:

20 CQECLRLKK
 S R

for which four sets of degenerate oligos were made. Namely:

25 ZnFA3' 5' TTC TTI AGI CGG CAC TCT TGG CA 3'
 T A T C A

ZnFB3' 5' TTC TTI AAI CGG CAC TCT TGG CA 3'
 T A T C A

30 ZnFC3' 5' TTC TTI AGI CTG CAC TCT TGG CA 3'
 T A T C A

ZnFD3' 5' TTC TTI AAI CTG CAC TCT TGG CA 3'
 T A T C A

35 The PCR amplification was carried out using a randomly primed cDNA library made from mRNA isolated from 4th and 5th instar *Heliothis virescens* larvae. The amplification

- was performed using 10^8 pfus (plaque forming units) in 50mM KCl, 20mM Tris HCl pH 8.4, 15mM MgCl₂, 200mM dNTPs (an equimolar mixture of dCTP, dATP, dGTP and dTTP), 100ng of ZnFAS' and ZnF3' mixture. The conditions used in the reaction followed the hot start protocol whereby the reaction mixture was heated to 94°C for 5 minutes after which 1
- 5 U of Taq polymerase was added and the reaction allowed to continue for 35 cycles of 93°C for 50 seconds, 40°C for 1 minute and 73°C for 1 minute 30 seconds. The PCR products were fractionated on a 2%(w/v) agarose gel and the fragment migrating between 100 and 200bp markers was isolated and subcloned into the vector pCRII (Invitrogen). The sequence of the insert was determined using Sequenase (USB).
- 10 The resulting sequence was translated and a database search carried out. The search recovered sequences matching to the DNA binding domain of the *Drosophila* ecdysone receptor, retinoic acid receptor and the thyroid receptor. Thus, the sequence of the insert in this plasmid, designated pCRIIZnf, is a *Heliothis* ecdysone cognate sequence (Figure 1) and was used to screen a cDNA library in order to isolate the complete open reading frame.

15

B. Library screening

The randomly primed cDNA 4th/5th Instar *Heliothis virescens* library was plated and replicate filter made from the plates. The number of plaques plated was 500,000. The insert fragment of pCRIIZnf was reamplified and 50ng were end labelled using T4 Polynucleotide Kinase (as described in Sambrook et al 1990).

20 The filter were prehybridised using 0.25%(w/v) Marvel, 5 X SSPE and 0.1%(w/v) SDS at 42°C for 4 hours. The solution in the filters was then replaced with fresh solution and the denatured probe added. The hybridisation was carried out overnight at 42°C after which the filter were washed in 6 X SSC + 0.1%(w/v) SDS at 42°C followed by another wash at 25 55°C. The filter were exposed to X-ray film (Kodak) for 48 hours before processing.

The developed film indicated the presence of one strong positive signal which was plaque purified and further characterised. The lambda ZAP II phage was in vivo excised (see Stratagene Manual) and the sequence determined of the resulting plasmid DNA. The clone known as pSK19R (or 19R) contained a 1.933kb cDNA fragment with an open reading frame of 467 amino acids (Figure 2). pSK19R was deposited with the NCIMB on 20 June 1995 and has been accorded the deposit No NCIMB 40743.

30 Further analysis of pSK19R revealed that a 340 bp EcoRI fragment mapping at the 5' end of pSK19R has strong and significant similarities to a *Drosophila* cDNA encoding glyceraldehyde-3-phosphate dehydrogenase. In order to isolate the correct 5'end sequence belonging to *Heliothis*, the random primed library was re-screened using a probe containing the 5'end of the pSK19R belonging to *Heliothis* ecdysone receptor. The probe was made by PCR using the sense oligonucleotide HecRH3C (5' aattaagcttccaccatgccgttaccaatgccaccgaca

3') and antisense oligonucleotide HecrNdeI (5' cttcaaccgacactctgac 3'). The PCR was carried out as described by Hirst et al., 1992) where the amount of radioisotope used in the labelling was 50uCi of a ³²P-dCTP and the PCR was cycled for 1 minute at 94°C, 1 minute at 60°C and 1 minute at 72°C for 19 cycles. The resulting 353bp radio labelled DNA fragment was denatured and added to prehybridised filters as described for the isolation of pSK19R.

The library filters were made from 15 plates each containing 50000 pfus. The library filters were hybridised at 65°C and washed in 3XSSPE + 0.1%SDS at 65°C twice for 30 minutes each. The filters were further washed with 1XSSPE + 0.1%SDS for 30 minutes and exposed to X-ray film (Kodak) overnight. The film was developed and 16 putative positive plaques were picked. The plaques were re-plated and hybridised under the exact same conditions as the primary screen resulting in only one strong positive. The strong positive was consistently recognised by the probe and was plaque purified and *in vivo* excised. The resulting plasmid pSK16.1 was sequenced (Seq 1D3) which revealed that the 5' end of the clone extended by 205 bp and at the 3' end by 653 bp and resulting in a DNA insert of 2.5 kb. Conceptual translation of the 205 bp yielded 73 amino acids with high similarity to the *Drosophila*, *Aedes aegypti*, *Manduca* and *Bombyx* sequences of the ecdysone receptor B1 isoform. However, the whole of the 5' end sequence is not complete since a Methionine start site was not found with a stop codon in frame 5' of the methionine. In order to isolate the remainder of the 5' end coding sequences a 5'RACE protocol (Rapid Amplification of cDNA Ends) was carried out using the BRL-GIBCO 5'RACE Kit. Two types of cDNA were synthesised where the first one used a specific oligonucleotide :

16PCR2A 5' cagctccaggcccgcgtcg3'

and the second type used random hexamers (oligonucleotide containing 6 random nucleotides). Each cDNA was PCR amplified using the oligonucleotides anchor primer :

25 BRL-GIBCO 5' cuacuacuacuaggccacgcgtcgactagtacgggiigggiiigggiiig 3'

and 16PCR2A and cycled for 1 minute at 94°C, 1 minute at 60°C and 1 minute at 72°C for 35 cycles. The reaction conditions were 20mM Tris-HCl (pH8.4), 50mM KCl, 1.5mM MgCl₂, 400nM of each anchor and 16PCR2A primers, 200mM dNTPs (dATP,dCTP,dGTP and dTTP) and 0.02 U/ml *Taq* DNA polymerase. Dilutions of 1:50 of the first PCR reactions were made and 1ml was used in a second PCR with oligonucleotides UAP :

(Universal Amplification Primer 5' caucaucauggccacgcgtcgactagtac 3')

and 16RACE2 :

(5' acgtcacccatcgagcgtctccattc 3').

The conditions and cycling were the same as those followed for the first PCR.

35 Samples of each PCR were run and a Southern blot carried out which was probed with a 5' specific primer :

(16PCR1 5' cgctggataacaacggaccattc 3').

This primer is specific for the 5' most sequence of pSK16.1 and was hybridised at 55°C using the standard hybridisation buffer. The filter was washed at 55°C 3 times in 3XSSPE + 0.1%SDS and exposed to X-ray film for up to 6 hours. The developed film revealed bands recognised by the oligonucleotide migrating at 100bp and 500bp (relative to the markers). A sample of the PCR reaction (4 in total) was cloned into the pCRII vector in the TA cloning kit (Invitrogen). Analysis of 15 clones from 4 independent PCRs yielded sequence upstream of pSK16.1 (Figure 4).

Translation of the ORF results in a 575 amino acid protein with high similarity in the DNA and ligand binding domains when compared to the ecdysone receptor sequences of 10 *Drosophila*, *Aedes aegypti*, *Chironomus tentans*, *Manduca sexta* and *Bombyx mori* (Figure 5). Interestingly, the N-terminal end of the *Heliothis* sequence has an in frame methionine start which is 20 amino acids longer than that reported for *Drosophila*, *Aedes aegypti* and *Manduca sexta*. However, the extended N-terminal end in the *Heliothis* EcR does not have similarity to that of *Bombyx mori*. Finally, the C-terminal end of the different B1 isoform 15 ecdysone receptor sequences diverge and do not have significant similarity.

C. Northern Blot Analysis

The sequence identified by screening the library is expected to be expressed in tissues undergoing developmental changes, thus mRNA from different developmental stages of *H. virescens* were isolated and a norther blot produced. The mRNAs were isolated from eggs, 1st, 2nd, 3rd, 4th and 5th instar larvae, pupae and adults. The northern blot was hybridised with a NdeI/XhoI DNA fragment from pSK19R encompassing the 3'end of the DNA binding domain through to the end of the ligand binding domain. The hybridisation was carried out in 1%(w/v)Marvel, 5X SSPE, 0.1%(w/v) SDS at 65°C for 18 to 24 hours. 20 25 The filters were washed in 3X SSPE + 0.1%(w/v) SDS and 1X SSPE + 0.1%(w/v) SDS at 65°C. The filter was blotted dry and exposed for one to seven days. The gene recognises two transcripts (6.0 and 6.5 kb) which appear to be expressed in all stages examined; however, the levels of expression differ in different stages. It should be noted that the same two transcripts are recognised by probes specific to the DNA binding domain and the ligand binding domain, 30 indicating that the two transcripts arise from the same gene either by alternative splicing or alternative use of polyadenylation sites.

In summary, adult and 5th instar larvae have lower levels of expression while all other tissues have substantial levels of expression.

Example II Expression of *Heliothis* ecdysone receptor in Mammalian cells

5 To demonstrate that the cDNA encodes a functional ecdysone receptor, effector constructs were generated containing the HEcR under the control of the CMV (cytomegalovirus) promoter, and the DNA expressed in mammalian cells.

Effector constructs

A first mammalian expression plasmid was constructed by placing a HindIII/NotI pSK19R fragment into the pcDNA3 HindIII/NotI vector resulting in pcDNA319R (Figure 7).

10 A second effector plasmid was constructed wherein the non-coding region of the cDNA 19R was deleted and a consensus Kozak sequence introduced. The mutagenesis was carried out by PCR amplifying a DNA fragment with the oligo HecRH3C :

5'at_attaagcttccaccatgccgttaccaatgccaccgaca 3'

containing a unique HindIII restriction enzyme recognition site followed by the mammalian 15 Kozak consensus sequence, and HecRNdeI :

5'cttcaaccgacactcctgac 3'.

The resulting 353bp PCR fragment was restriction enzyme digested with HindIII and NdeI, gel purified and ligated with 19R NdeI/NotI fragment into a pcDNA3 HindIII/NotI vector resulting in pcDNA3HecR.

20 A third effector construct was made with the 5' end sequences of pSK16.1 by PCR. The PCR approach involved PCR amplifying the 5' end sequences using a 5' oligonucleotide containing a HindIII restriction cloning site, the Kozak consensus sequence followed by nucleotide sequence encoding for a Methionine start and two Arginines to be added to the 5' end of the amplified fragment :

25 (16H3K 5' at_attaagcttgccgc_catgcggcgacgctggataacaacggaccattc 3'), the 3' oligonucleotide used was HecrNdeI. The resulting fragment was restriction enzyme digested, gel purified and subcloned with an NdeI/NotI 19R fragment into pcDNA3 NdeI/NotI vector. The plasmid was named pcDNA3H3KHEcR.

30 A fourth effector construct was produced which contains the extended N-terminal end sequence obtained from the 5'RACE experiment. Thus, a PCR approach was followed to introduce the new 5' end sequences in addition to a consensus Kozak sequence and a HindIII unique cloning sequence. The sense oligonucleotide used was RACEH3K :

5' at_attaagcttgccgc_catgtccctggcgctcgat 3',

35 while the antisense primer was the same as that used before (HecrNdeI). The cloning strategy was the same as used for the pcDNA3H3KHEcR to give rise to pcDNA3RACEH3KHEcR.

The PCR mutagenesis reactions were carried out in the same manner for all constructs. The PCR conditions used were 1 minute at 94°C, 1 minute at 60°C and 1 minute

at 72°C for 15 cycles. The reactions conditions were 50mM Tris-HCl (pH8.4), 25mM KCl, 200mM dNTPs (dATP, dCTP, dGTP and dTTP), 200nM of each oligonucleotide and 2.5U/Reaction of *Taq* DNA polymerase. For each construct at least 5 independant PCR reactions were carried out and several clones were sequenced to insure that at least one is mutation free.

5 **Reporter construct**

The reporter plasmid to be co-transfected with the expression vector contained 4 copies of the Hsp27 ecdysone response element (Riddihough and Pelham, 1987) fused to B-globin promoter and the B-Galactosidase gene. The tandem repeats of the ecdysone response 10 element were synthesised as two complementary oligonucleotides which when annealed produced a double stranded DNA molecule flanked by an SpeI site at the 5' end and a Clal site at the 3' end :

Recr3A

15 5'ctagtagacaagggttcaatgcacttgtccaataagcttagacaagggttcaatgcacttgtccaatgaattcagacaagggttcaat gcacttgtccaatctgcagagacaagggttcaatgcacttgtccaatat 3'

Recr3B

5'cgatattggacaagtgcattgaacccttgtctgcagattggacaagtgcattgaacccttgtctgaattcattggacaagtgcattg aacccttgtctaagcttattggacaagtgcattgaacccttgtcta 3'.

The resulting 135bp DNA fragment was ligated to the vector pSWBGAL SpeI/Clal 20 resulting in pSWREcR4 (Figure 8). The co-transfection of the two plasmid should result in B-galactosidase activity in the presence of ligand. The experiment relies upon the presence of RXR (a homologue of ultraspiracle) in mammalian cells for the formation of an active ecdysone receptor.

25 **Mammalian transfection methods**

Transfections of mammalian cell lines (CHO-K1 Chinese hamster ovary)- ATCC number CCL61 or cos-1 (Monkey cell line) were performed using either calcium phosphate precipitation (Gorman, Chapter 6 of "DNA cloning: a practical approach. Vol 2 D.M. Glover ed.(1985) IRL Press, Oxford) or using LipofectAMINE (Gibco BRL Cat. No. 18324-012, following manufacturers instructions). Human Epithelial Kidney-293 cells were transfected 30 using analogous methods.

Results - Native HEcR drives transient reporter gene expression in mammalian cells

Co-transfection of pcDNA3H3KHEcR (Effector) and reporter constructs into Human Epithelial Kidney 293 cells (HEK293) in the presence of either Muristerone A or RH5992 resulted in a 2-3 fold induction of reporter activity compared to the no chemical controls 35 (Figure 9). The HEK293 cells were used since they are known to have constitutive levels of α RXR which have been demonstrated to be necessary for *Drosophila* EcR activation by Muristerone A (Yao., et al., 1993). Moreover, to further investigate the need for RXR

interactions, a α RXR was co-transfected into HEK293 cells (along with the effector and reporter) resulting in a 9 fold induction of reporter activity compared to the untreated cells (Figure 9). The co-transfection of α RXR with reporter and effector increased by four fold the reporter activity compared to cells transfected with effector and reporter alone. Induction was observed both in the presence of either Muristerone A or RH5992. These data clearly demonstrate that the cDNA HEcR encodes a functional ecdysone receptor. Moreover, The ability of HEcR to complex with α RXR and bind Muristerone A or RH5992 provide evidence for the usage of the entire HEcR as a component of a mammalian gene switch. In particular, it offers the advantage of reducing uninduced expression of target gene since ecdysone receptor and response elements are not present in mammalian cells.

Example III - Chimeric constructs and ligand validation in Maize Protoplasts

In order to apply the ecdysone receptor as an inducible system it was deemed necessary to simplify the requirements of the system by avoiding the need of a heterodimer formation to obtain an active complex. The glucocorticoid receptor is known to form homodimers and chimeric constructs of the glucocorticoid receptor transactivating and DNA binding domains fused to the ecdysone receptor hinge and ligand binding domains have been shown to be active as homodimers in mammalian cells in the presence of Muristerone A (an ecdysone agonist)(Christopherson et al., 1992). However, the chimeric receptor is not responsive to 20-hydroxyecdysone (Christopherson et al., 1992).

The analysis of the activation of the glucocorticoid/*Heliothis* ecdysone chimeric receptor entailed the production of two other control effector constructs. The first one of the constructs contained the intact glucocorticoid receptor while the second one contained a glucocorticoid/*Drosophila* ecdysone chimeric receptor.

Effector constructs

(i) Glucocorticoid receptor Maize expression construct

The glucocorticoid receptor DNA for the Maize transient expression construct was produced via the polymerase chain reaction (PCR) of Human Fibrosarcoma cDNA (HT1080 cell line, ATCC#CC1121) library (Clontech)(see Hollenberg et al., 1985). The PCR approach taken was to amplify the 2.7kb fragment encoding the glucocorticoid receptor in two segments. The first segment entails the N-terminal end up to and including the DNA binding domain while the second fragment begins with the hinge region (amino acid 500) thought to the end of the reading frame. Thus, the PCR primer for the N-terminal end segment was designed to contain an EcoRI site and the Kozak consensus sequence for translation initiation

GREcoRI 5'atgaaattccaccatggactccaaagaatcattaactc 3'.

The 3'end primer contains a XhoI site in frame with the reading frame at amino acid 500 of the published sequence :

GRXhoI 5' gagactcctgttagtggcctcgagcattcc~~ttt~~tat~~ttt~~tttc 3'.

5 The second fragment of the glucocorticoid receptor was produced with a 5' end oligonucleotide containing an XhoI site in frame with the open reading frame at the begining of the hinge region (amino acid 500) :

GRHinge 5' attctcggaggattcagcaggccactacaggag 3'

while the 3' end oligonucleotide contained an EcoRI site 400 bp after the stop codon :

GRStop 5' attgaattcaatgctatcgtaactatacagg 3'.

10 The glucocorticoid receptor PCR was carried out using Vent polymerase (Biolabs) under hot start conditions followed by 15 cycles of denaturing (94°C for 1 minute), annealing (66°C for 1 minute) and DNA synthesis (72°C for 3 minute). The template was produced by making first strand cDNA as described in the TA cloning kit (Invitrogen) after which the PCR was carried out in 10mM KCl, 10mM (NH₄)₂SO₄, 20mM TRIS-HCl pH 8.8, 2 mM MgSO₄, 15 0.1% (v/v) Triton X-100, 200 mM dNTPs, 100ng of each Primer and 2 U of Vent Polymerase. The PCR products was restriction enzyme digested with EcoRI and XhoI and subcloned into pBluescript SK (pSK) EcoRI. The resulting plasmid pSKHGI was sequenced and found to lack any mutations from the published sequences (apart from those introduced in the PCR primers) (Hollenberg et al., 1985).

20 The 2.7kb EcoRI fragment was subcloned into the vector pMF6PAT EcoRI resulting in pMF6HGIPAT (Figure 10).

(ii) Maize expression construct containing a Glucocorticoid/ *Drosophila* ecdysone chimeric receptor.

25 The glucocorticoid receptor portion of the chimeric receptor was isolated from pSKHGI by producing a 1.5kb BamHI/XhoI restriction fragment containing the N-terminal end up to and including the DNA binding domain.

30 The *Drosophila* ecdysone receptor portion was isolated through PCR of first stand cDNA prepared from *Drosophila* adult mRNA. The PCR was carried out using a 5' oligonucleotide containing a Sall site (ie. *Drosophila* ecdysone receptor contains a XhoI site at the end of the ligand binding domain) which starts at the begining of the hinge region : amino acid 330, Ecr8 attgtcgacaacggccggaatggtcgtccggag 3'.

The 3' end oligonucleotide contains an BamHI site adjacent to the stop codon : EcRstop 5' tcgggctttgttaggatcctaagccgtggtcgaatgctccacttaac 3'.

35 The PCR was carried out under the conditions described for the amplification of the Glucocorticoid receptor and yielded a 1.6 kb fragment. The fragment was introduced into

pSK SalI/BamHI and the sequence determined and compared to the published one (Koelle et al., 1991).

The maize transient expression plasmid was produced by introducing into pMF6 BamHI vector the 1.5kb BamHI/XhoI glucocorticoid receptor fragment and the 1.6kb 5 SalI/BamHI *Drosophila* receptor portion to yield the chimeric plasmid pMF6GREcRS (Figure 9).

(iii) Construction of the Glucocorticoid/*Heliothis* ecdysone chimeric receptor Maize transient expression plasmid.

10 The Glucocorticoid receptor portion of the chimera was produced as described in Example II(ii). The production of the *Heliothis* ecdysone receptor portion involves the introduction of a SalI recognition site at the DNA binding/hinge domain junction (amino acid 229). The addition of the SalI site :

Hecrsal 5'attgtcgacaaaggccccgagtgcgtggtgccggag 3'

15 was achieved via PCR mutagenesis making use of an unique AccI site 107bp downstream of the junction point (or 1007 bp relative to Seq 1D No 4):

Hecracc 5' tcacattgcatgtggaggcatg 3'.

20 The PCR was carried out using *Taq* polymerase (2.5 U) in a reaction buffer containing 100ng of template DNA (pSK19R), 100ng of Hecrsal and Hecracc, 20 mM TRIS-HCl pH 8.4, 50mM KCl, 10mM MgCl₂, 200mM dNTPs. The reaction was carried out with an initial denaturation of 3 minutes followed by 15 cycles of denaturation (1 minute at 94°C), annealing (1 minute at 60°C) and DNA synthesis (1 minute at 72°C). The DNA was restriction enzyme digested and subcloned into pSK SalI/SacI with the 1.2kb AccI/SacI 3' end HecR fragment to yield pSK HeCRDEF (or containing the hinge and ligand binding domains of the *Heliothis* 25 ecdysone receptor). The construction of the maize transient expression plasmid containing the Glucocorticoid/*Heliothis* ecdysone chimeric receptor involved the ligation of pMF6 EcoRI/SacI with the 1.5kb EcoRI/XhoI fragment of Glucocorticoid receptor N-terminal end and the 1.2 kb SalI/SacI fragment of pSk HEcRDEF to yield pMF6GRHEcR (Figure 10).

Reporter plasmids

30 Two reporter plasmids were made by inserting the into p221.9 or p221.10 BamHI/HindIII vectors two pairs of oligonucleotides containing six copies of the glucocorticoid response element (GRE). The two sets of oligonucleotides were designed with restriction enzyme recognition sites so as to ensure insertion of the two pairs in the right orientation. The first oligonucleotide pair GRE1A/B is 82 nucleotides long and when annealed 35 result in a DNA fragment flanked with a HindIII site at the 5' end and a SalI site at the 3' end : GRE1A

5'agcttcgactgtacaggatgttctagctactcgagtagctagaacatcctgtacagtgcgagtagctagaacatcctgtacag 3'

GRE1B

5'tcgactgtacaggatgttctactcgactgtacaggatgttctagctactcgagtcgctagaacatcctgta cagtcta 3'.

The second pair of oligonucleotides is flanked by a SalI site at the 5' end and a BamHI site at the the 3' end

5 GRE2A 5' tcgacttagctagaacatcctgtacagtgcgaggtagctagaacatcctgt
acagtcgaggtagctagaacatcctgtacag 3'

GRE2B 5'gatcctgtacaggatgttctactcgactgtacaggatgttctagctactcgactgtacaggatgttctagctag 3'.

The resulting plasmids were named p221.9GRE6 (Figure 13) and p221.10GRE6 (Figure 14)(used in later Example). The difference between p221.9 and p221.10 plasmids is 10 that p221.9 contains the -60 35SCaMV minimal promotor while p221.10 (p221.10GRE6) contains the -46 35SCaMV minimal promotor.

Method

Protoplasts were isolated from a maize suspension culture derived from BE70 x A188 embryogenic callus material, which was maintained by subculturing twice weekly in MS0.5_{mod}. 15 (MS medium supplemented with 3% sucrose, 690mg/l proline, 1g/l myo-inositol, 0.2g/l casein acid hydrolysate, 0.5mg/l 2,4-D, pH5.6). Cells from suspensions two days post subculture were digested in enzyme mixture (2.0% Cellulase RS, 0.2% Pectolyase Y23, 0.5M Mannitol, 5mM CaCl₂H₂O, 0.5% MES, pH5.6, ~660mmol/kg) using ~10ml/g cells, incubating at 25°C, dim light, rotating gently for ~2 hours. The digestion mixture was sieved sequentially through 20 250µm and 38µm sieves, and the filtrate centrifuged at 700rpm for 3.5 minutes, discarding the supernatant. The protoplasts were resuspended in wash buffer (0.358M KCl, 1.0mM NH₄NO₃, 5.0mM CaCl₂H₂O, 0.5mM KH₂PO₄, pH4.8, ~670mmol/kg) and pelleted as before. This washing step was repeated. The pellet was resuspended in wash buffer and the protoplasts were counted. Transformation was achieved using a Polyethylene glycol method 25 based on Negruiti et.al. Protoplasts were resuspended at 2 x 10⁶/ml in MaMg medium (0.4M Mannitol, 15mM MgCl₂, 0.1% MES, pH5.6, ~450mmol/kg) aliquotting 0.5ml / treatment (i.e. 1x10⁶ protoplasts / treatment). Samples were heat shocked at 45°C for 5 minutes then cooled to room temperature. 10µg each of p221.9GRE6 and pMF6HR1PAT (GR) (1mg/ml) / treatment were added and mixed in gently, followed by immediate addition of 0.5ml warm 30 (~45°C) PEG solution (40% PEG 3,350MW in 0.4M Mannitol, 0.1M Ca(NO₃)₂, pH8.0), which was mixed in thoroughly but gently. Treatments were incubated at room temperature for 20-25 minutes, then 5ml 0.292M KCl (pH5.6, ~530mmol/kg) was added step-wise, 1ml at a time, with mixing. The treatments were incubated for a further 10-15 minutes prior to pelleting the protoplasts by centrifuging as before. Each protoplast treatment was 35 resuspended in 1.5ml culture medium (MS medium, 2% sucrose, 2mg/l 2,4-D, 9% Mannitol, pH5.6, ~700mmol/kg) +/- 0.0001M dexamethasone (glucocorticoid). The samples were incubated in 3cm dishes at 25°C, dark, for 24-48 hours prior to harvesting. Fluorometric

assays for GUS activity were performed with the substrate 4-methylumbelliferyl-D-glucuronide using a Perkin-Elmer LS-35 fluorometer (Jefferson et al., 1987). Protein concentration of tissue homogenates were determined by the Bio-Rad protein assay (Bradford, 1976). The method was repeated for each effector construct.

5 Results

Reporter gene assay

A reporter gene construct (p221.9GRE6) was generated containing the GUS reporter gene under the control of a -60 CaMV 35S promoter with 6 copies of the glucocorticoid response element. To test this construct was functional in maize protoplasts a co-transformation assay was performed with the reporter construct p221.9GRE6 and the effector construct pMF6HR1PAT (GR) construct containing the entire glucocorticoid receptor.

10 Figure 15 shows that Reporter p221.9GRE6 alone or reporter plus effector pMF6HR1PAT (GR) with no activating chemical gave no significant expression. When reporter plus effector were co-transformed into maize protoplasts in the presence of 0.0001M dexamethasone (glucocorticoid), a significant elevation of marker gene activity was observed (Figure 15). The response is specific to glucocorticoid as the steroid Muristerone A does not lead to induced levels of expression. These studies clearly show the reporter gene construct p221.9GRE6 is capable of monitoring effector /ligand mediated gene expression.

15 Chimeric ecdysone effector constructs mediate inducible expression in maize transient
20 protoplasts assays

A chimeric effector plasmid pMF6GREcRS was constructed, containing the ligand binding domain from the *Drosophila* ecdysone receptor and the DNA binding and transactivation domain from the glucocorticoid receptor. To confirm the reporter gene construct p221.9GRE6 could respond to a chimeric ecdysone effector construct, a series of co-transformation into maize protoplasts was performed.....

25 Figure 16 shows that reporter (p221.9GRE6) alone or reporter plus effector (pMF6GREcRS) with no activating chemical, gave no significant expression in maize protoplasts. When reporter plus effector were co-transformed into maize protoplasts in the presence of 100 μ M Muristerone A, a significant elevation of marker gene activity was observed. The response was specific to Muristerone A, as the steroid dexamethasone did not lead to induced levels of expression. These studies clearly showed the reporter gene construct p221.9GRE6 is capable of monitoring chimeric ecdysone effector /ligand mediated gene expression.

30 A second chimeric effector construct pMF6GRHEcR, was generated containing the ligand binding domain from *Heliothis* ecdysone receptor. When co-transformed into maize protoplasts with the reporter plasmid p221.9GRE6, no response to 100 μ M Muristerone or

100 μ M dexamethasone was observed (Figure 17). These data clearly show the *Drosophila* and *Heliothis* ligand binding domains exhibit different properties.

When the effector plasmid pMF6GREcRS, containing the ligand binding domain from *Drosophila*, was tested with the reporter p221.9GRE6 in presence of the non-steroidal 5 ecdysone agonists RH5849 and RH5992 (mimic), no chemical induced reporter gene activity was observed (Figures 18 and 19).

When the effector plasmid pMF6GRHEcR, containing the ligand binding domain from *Heliothis*, was tested with the reporter p221.9GRE6 in presence of the non-steroidal 10 ecdysone agonists RH5992 (mimic), significant chemical induced reporter gene activity was observed (Figure 20). These data demonstrate the ligand binding domain from *Heliothis* has different properties to the *Drosophila* receptor in that the former responded to the non-steroidal ecdysteroid agonist RH5992. Figure 21 demonstrates the effector plasmid pMF6GRHEcR confers RH5992 dependant inducibility on the reporter p221.9GRE6 in a dose responsive manner. Induction was observed in a range from 1 μ M-100 μ M RH5992.

15

Example IV - Testing of effector vectors in Tobacco protoplasts

The experiments carried out in the previous example demonstrated the specific effect of RH5992 (mimic) on pMF6GRHEcR in maize protoplasts. It is the aim in this example to 20 show the generic application to plants of the glucocorticoid/*Heliothis* ecdysone chimeric receptor switch system. Tobacco shoot cultures cv. Samsun, were maintained on solidified MS medium + 3% sucrose in a controlled environment room (16 hour day / 8 hour night at 25°C, 55% R.H), were used as the source material for protoplasts. Leaves were sliced parallel to the mid-rib, discarding any large veins and the slices were placed in CPW13M 13% 25 Mannitol, pH5.6, ~860mmol/kg) for ~1 hour to pre-plasmolyse the cells. This solution was replaced with enzyme mixture (0.2% Cellulase R10, 0.05% Macerozyme R10 in CPW9M (CPW13M but 9% Mannitol), pH5.6, ~600mmol/kg) and incubated in the dark at 25°C overnight (~16 hours). Following digestion, the tissue was teased apart with forceps and any large undigested pieces were discarded. The enzyme mixture was passed through a 75 μ m 30 sieve and the filtrate was centrifuged at 600rpm for 3.5 minutes, discarding the supernatant. The pellet was resuspended in 0.6M sucrose solution and centrifuged at 600rpm for 10 minutes. The floating layer of protoplasts was removed using a pasteur pipette and diluted with CPW9M (pH5.6, ~560mmol/kg). The protoplasts were again pelleted by centrifuging at 600rpm for 3.5 minutes, resuspended in CPW9M and counted. A modified version of the 35 PEG-mediated transformation above was carried out. Protoplasts were resuspended at 2x10⁶/ml in MaMg medium and aliquotted using 200 μ l / treatment (i.e. 4x10⁵ protoplasts / treatment). 20 μ g each of pMF6GRHEcRS and p221.9GRE6 DNA (1mg/ml) were added

followed by 200 μ l PEG solution and the solutions gently mixed. The protoplasts were left to incubate at room temperature for 10 minutes before addition of 5ml MSP19M medium (MS medium, 3% sucrose, 9% Mannitol, 2mg/l NAA, 0.5mg/l BAP, pH5.6, ~700mmol/kg) +/- 10 μ M RH5992. Following gentle mixing, the protoplasts were cultured in their tubes, lying horizontally at 25°C, light. The protoplasts were harvested for the GUS assay after ~24 hours.

5 Effector construct

(i) Construction of a Dicotyledonous expression vector

The vector produced is a derivative of pMF6. pMF6GREcRS was restriction enzyme digested with PstI to produce 3 fragments namely, 3.4(Adh Intronless pMF6), 3.2(GREcRS) 10 and 0.5(Adh intron I) kb). Isolation and religation of the 3.4 and 3.2 kb fragments resulted in pMF7GREcRS (Figure 22). pMF7GREcRS was restriction enzyme digested with EcoRI/SacI resulting in the 3.4kb pMF7 EcoRI/SacI vector which when isolated and purified was ligated to a 1.5 kb EcoRI/XhoI N-terminal end of the glucocorticoid receptor and the 1.2 kb SalI/SacI *Heliothis* ecdysone C-terminal end sequences to produce pMF7GRHEcR (Figure 15 23).

Reporter plasmid

The reporter plasmids constructed for the maize transient experiments were the same as those used without alteration in the tobacco leaf protoplast transient expression experiments.

20 Results - Chimeric ecdysone effector constructs mediate inducible expression in tobacco transient protoplast assays

Experiments were performed to demonstrate that the effector plasmid pMF7GRHEcR can confer chemical dependant inducible expression on the reporter p221.9GRE6 in tobacco mesophyll protoplasts.

25 Figure 24 shows that reporter (p221.9GRE6) alone or reporter plus effector (pMF7GRHEcR) with no activating chemical, gave no significant expression in tobacco protoplasts. When reporter plus effector were co-transformed into tobacco protoplasts in the presence of 10 μ M RH5992, a significant elevation of marker gene activity was observed. These data show a chimeric ecdysone effector construct, containing the *Heliothis* ligand 30 binding domain can confer non-steroidal ecdysteroid dependant expression on reporter gene constructs in both monocotyledonous and dicotyledonous species.

Example V - Chimeric activity in Mammalian cells**Effector constructs****5 (i) Construction of Glucocorticoid/*Heliothis* ecdysone chimeric receptor.**

The mammalian expression vector used in this experiment was pcDNA3 (Invitrogen). The GRHEcR 2.7kb BamHI DNA fragment (isolated from pMF6GRHEcR) was introduced into the pcDNA3 BamHI vector. The recombinants were oriented by restriction enzyme mapping. The DNA sequence of the junctions was determined to ensure correct orientation and insertion (pcDNA3GRHEcR, Figure 25).

Reporter construct

10 The reporter plasmid for mammalian cell system was produced by taking pSWBGAL plasmid and replacing the CRESW SpeI/Clal fragment for a synthetic 105 bp DNA fragment containing 4 copies of the glucocorticoid response element (GRE) and flanked by SpeI at the 15 5' end and AflII at the 3' end.

The oligonucleotides were synthesised using the sequences :

GREspeI

5'ctagtgtacaggatgttctactcgagtagctagaacatccgtacagtcgagtagctagaacatccgtacagtcgagtagct
agaacatccgtacac 3'

20 GREafl2

5'ttaagtgtacaggatgttctactcgactgtacaggatgttctactcgactgtacaggatgttctactcgagtagct
gaacatccgtacaa 3'.

The two oligonucleotides were purified annealed and ligated to pSWBGAL SpeI/AflII to produce pSWGRE4 (Figure 26).

25 Results - Chimeric HEcR drives transient reporter gene expression in mammalian cells

No expression was detected when a reporter gene construct pSWGRE4, comprising of a minimal β-globin promoter containing four copies of the glucocorticoid response element, fused to a β-galactosidase reporter gene, was introduced into CHO cells. Similarly, no expression was detected when pSWGRE4 and an effector plasmid pCDNA3GRHEcR, 30 containing the transactivation and DNA binding domain from the glucocorticoid receptor and the ligand binding domain from the *Heliothis* ecdysone receptor, under the control of the CMV promoter were co-transformed into CHO-K1 or HEK293 cells. When co-transformed CHO (Figure 27) and HEK293 cells (Figure 28) were incubated in the presence of the non-steroidal ecdysone agonists RH5992 (mimic), significant chemical induced reporter gene 35 activity was observed. Equally, induction of reporter activity was observed when HEK293 cells transfected with pcDNA3GRHEcR and reporter were treated with Muristerone A (Figure 28).

Example VI - Screening system allows new chemical activators and modified ligand binding domains to be tested in Mammalian cells

5 The basis of a screening system are in place after the demonstration that the chimeric receptor was activated in the presence of RH5992. A screen was carried out using CHO cells transiently transfected with both pSWGRE4 (reporter) and pcDNA3GRHEcR (effector) constructs. In the first instance 20 derivatives compounds of RH5992 were screened. It was observed that 7 out of the 20 compounds gave an increased reporter gene activity compared to untreated cells. A second screen was carried out in which 1000 randomly selected compounds were applied to transiently transfected CHO cells. Two compounds were found to activate reporter gene activity above that from the untreated controls. The second screen suggest that this cell based assay is a robust and rapid way to screen a small library of compounds, where a thousand compounds can be put through per week.

15

Example V - Stably transformed Tobacco plants

Stable Tobacco vectors

20 The components of the stable Tobacco vectors were put together in pBluescript prior to transfer into the binary vector. The production of stable transformed plants entails the production of a vector in which both components of the switch system (ie. effector and reporter) are placed in the same construct to then introduce into plants.

The methodology described below was used to produce four different stable Tobacco vectors. The method involves three steps:

25

1. pBluescript SK HindIII/EcoRI vector was ligated to either GRE6-4635SCaMVGUSNOS HindIII/EcoRI (from p221.10GRE6) or GRE6-6035SCaMVGUSNOS HindIII/EcoRI (from p221.9GRE6) resulting in plasmid pSK-46 and pSK-60.

30

2. This step involves the addition of the chimeric receptor (35SGRHEcRNOS or 35SGRVP16HEcRNOS) to pSK-60 or pSK-46. Thus a pSK-60 (or pSK-46) XbaI vector was ligated with either the 3.4kb 35SGRHEcRNOS XbaI or the 3.0kb 35SGRVP16HEcRNOS XbaI DNA fragment to produce pSKES1 (pSKGRE6-6035SCaMVGUSNOS-35SGRHEcRNOS), pSKES2 (pSKGRE6-4635SCaMVGUSNOS-35SGRHEcRNOS), pSKES3 (pSKGRE6-6035SCaMVGUSNOS-35SGRVP16HEcRNOS) and pSKES4 (pSKGRE6-4635SCaMVGUSNOS-35SGRVP16HEcRNOS).

3. Transfer from pBluescript based vectors to binary vectors. The transfer of the ES1 (Figure 29) ES2 (Figure 30), ES3 (Figure 31) or ES4 (Figure 32) DNA fragments into the binary vector JR1 involves five steps:

5

(i) Restriction enzyme digestion of pSKES1 (ES2, ES3, and ES4) with ApaI and NotI to liberate the insert from the vector pBluescript.

(ii) The two DNA fragments were BamHI methylated for 2 hours at 37°C in TRIS-HCl, MgCl₂, 80uM SAM (S-adenosylmethionine) and 20 U of BamHI methylase.

10 (iii) Ligate a ApaI/NotI linker onto the fragment. The linker was designed to have an internal BamHI site :

ApaBNot1 5' cattggatcccttagc 3' and

ApaBNot2 5'ggccgctaaggatccaatgggcc 3'.

(iv) Restriction enzyme digest the protected and linked fragment with BamHI and 15 fractionate the products on a 1%(w/v) Agarose gel. The protected DNA fragment (5.5kb) was cut out of the gel and purified.

(v) A ligation of JRI BamHI vector with the protected band was carried out to produce JRIES1 (JRIES2, JRIES3 or JRIES4). The DNA of the recombinant was characterised by restriction mapping and the sequence of the junctions determined.

20 The plant transformation construct pES1, containing a chimeric ecdysone receptor and a reporter gene cassette, was transferred into Agrobacterium tumefaciens LBA4404 using the freeze/thaw method described by Holsters et al. (1978). Tobacco (*Nicotiana tabacum* cv Samsun) transformants were produced by the leaf disc method (Bevan, 1984). Shoots were regenerated on medium containing 100mg/l kanamycin. After rooting, plantlets were 25 transferred to the glasshouse and grown under 16 hour light/ 8 hour dark conditions.

Results - Chimeric ecdysone effector constructs mediate inducible expression in stably tobacco plants

Transgenic tobacco plants were treated in cell culture by adding 100μM RH5992 to 30 MS media. In addition seedlings were grown hydroponically in the presence or absence of RH5992. In further experiments 5mM RH5992 was applied in a foliar application to 8 week old glasshouse grown tobacco plants. In the three methods described uninduced levels of GUS activity were comparable to a wild type control, while RH5992 levels were significantly elevated.

Ecdysone switch modulation and optimisation**Example VIII - Yeast indicator strains for primary screen of chemical libraries**

5

A set of yeast indicator strains was produced to use as a primary screen to find chemicals which may be used in the gene switch. The properties of the desired chemicals should include high affinity resulting in high activation but with different physico-chemical characteristics so as to increase the scope of application of the technology. Moreover, the production of this strain also demonstrates the generic features of this switch system.

10 Effectuator vector

A base vector for yeast YCp15Gal-TEV-112 was generated containing:

Backbone - a modified version of pRS315 (Sikorski and Hieter (1989) Genetics 122, 19-27)- a shuttle vector with the LEU2 selectable marker for use in yeast;

15 ADH1 promoter (BamHI- Hind III fragment) and ADH1-terminator (Not I- Bam HI fragment) from pADNS (Colicelli et al PNAS 86, 3599-3603);

DNA binding domain of GAL4 (amino acids 1-147; GAL4 sequence is Laughon and Gesteland 19984) Mol. Cell Biol. 4, 260-267) from pSG424 (Sadowski and Ptashne (1989) Nuc. Acids Res. 17, 7539);

20 Activation domain - an acidic activation region corresponding to amino acids 1-107 of activation region B112 obtained from plasmid pB112 (Ruden et al (1991) Nature 350, 250-252).

The plasmid contains unique Eco RI, Nco I and Xba I sites between the DNA binding domain and activation domains.

25 Into this vector a PCR DNA fragment of the *Heliothis* ecdysone receptor containing the hinge, ligand binding domains and the C-terminal end was inserted. The 5' oligonucleotide is flanked by an NcoI restriction recognition site and begins at amino acid 259 :

HecriNcoI 5' aattccatggtacgacgacagttagacgtac 3'.

The 3' oligonucleotide is flanked by an XbaI site and encodes for up to amino acid

30 571:

HecRXbaI 5' ctgaggcttagagacggcggcggc 3'.

The PCR was carried out using vent polymerase with the conditions described in Example IA. The fragment was restriction enzyme digested with NcoI and XbaI purified and ligated into YCp15GALTEV112 NcoI/XbaI vector to produce YGALHeCRB112 or TEV-B112 (Figure 34). In order to reduce constitutive activity of the YGALHeCRB112 plasmid a YGALHeCR plasmid was produced in which the B112 activator was deleted by restriction enzyme digesting YGALHeCRB112 with XbaI/SpeI followed by ligation of the resulting

vector (ie. SpeI and XbaI sites when digested produce compatible ends)(TEV-8, Figure 33). An effector plasmid was constructed whereby the B112 transactivating domain was excised from YGalHecRB112 with XbaI and replaced with the VP16 transactivation domain DNA fragment (encoding amino acids 411 and 490 including the stop codon). The resulting vector 5 was named YGalHecRVP16 or TEVVP16-3 (Figure 35).

Reporter construction for yeast

The *S. cerevisiae* strain GGY1::171 (Gill and Ptashne (1987) Cell 51, 121-126), YT6::171 (Himmelfarb et al (1990) Cell 63, 1299-1309) both contain reporter plasmids consisting of the GAL4-responsive GAL1 promoter driving the *E. coli* B-galactosidase gene. 10 These plasmids are integrated at the URA3 locus. The reporter strain YT6::185 contains the reporter plasmid pJP185 (two synthetic GAL4 sites driving the B-galactosidase gene) integrated at the URA3 locus of YT6 (Himmelfarb et al). (Note- the parental strains YT6 and GGY1 have mutations in the GAL4 and GAL80 genes, so the reporter genes are inactive in the absence of any plasmids expressing GAL4 fusions).

15 Yeast assay

Standard transformation protocols (Lithium acetate procedure) and selection of colonies by growth of cells on selective media (leucine minus medium in the case of the YCp15Gal-TEV-112 plasmid)- as described in Guthrie and Fink (1991) Guide to Yeast Genetics and Molecular Biology: Methods in Enzymology Vol. 194 Academic Press) and the 20 reporter gene assay is a modification of that described in Ausabel et al (1993) Current Protocols in Molecular Biology (Wiley) Chapter 13).

Results - Automated screening system allows new chemical activators and modified ligand binding domains to be tested in yeast

An effector vector pYGALHEcRB112 has been generated containing a GAL4 DNA 25 binding domain, a B112 activation domain and the ligand binding region from *Heliothis virescens*. In combination with a GAL reporter vector, pYGALHEcRB112 form the basis of a rapid, high throughput assay which is cheap to run. This cell-based assay in yeast (*Saccharomyces cerevisiae*) will be used to screen for novel non-steroidal ecdysone agonists which may of commercial interest as novel insecticides or potent activators of the ecdysone 30 gene switch system. The demonstration of an efficient system to control gene expression in a chemical dependant manner, forms the basis of an inducible system for peptide production in yeast.

The yeast screening system forms the basis of a screen for enhanced ligand binding 35 using the lac Z reporter gene vector to quantitatively assay the contribution of mutation in the ligand binding domain. Alternatively, enhanced ligand binding capabilities or with a selection cassette where the lac Z reporter is replaced with a selectable marker such as uracil (URA 3), tryptophan (Trp1) or leucine (Leu2), and histidine (His). Constructs based on

pYGALHEcRB112 with alterations in the ligand binding domain are grown under selection conditions which impair growth of yeast containing the wild type ligand binding domain. Those surviving in the presence of inducer are retested and then sequenced to identify the mutation conferring resistance.

5

Example IX - Optimisation of chimeric receptor using a strong transactivator

Construction of mammalian expression plasmid with chimeric receptor containing Herpex Simplex VP16 protein sequences.

10 The construction of this chimeric receptor is based on replacing the sequences encoding for the glucocorticoid receptor transactivating domain with those belonging to the VP16 protein of Herpex simplex. Thus PCR was used to generate three fragments all to be assembled to produce the chimeric receptor. The PCRs were carried out as described in Example II, iii. The first fragment includes the Kozak sequences and methionine start site of 15 the glucocorticoid receptor to amino acid 152 of the glucocorticoid receptor. The oligonucleotides used for the generation of this fragment included an EcoRI site at the 5' end: GR1A 5' atatgaattccaccatggactccaaagaatc 3'
and at the 3' end a NheI restriction enzyme recognition site :
GR1B 5' atatgcttagctgtggggcagcagacacagcagtgg 3'.

20 The second fragment also belongs to the glucocorticoid receptor and begins with a NheI site in frame with amino acid 406 :
GR2A 5' atatgctagctccagctcctcaacagcaacaac 3'
and ends with a XhoI site at amino acid 500 :
GR2B 5' atatctcgagcaattcttttttttttc 3'.

25 The two fragments were introduced into pSKEcoRI/SacI in a ligation containing GR1A/B EcoRI/NheI, GR2A/B NheI/XhoI and HEcR SalI/SacI (from pSKHEcRDEF) to yield pSKGRDHEcR. The GR sequences and junctions of the ligation were found to be mutation free.

30 The third fragment to be amplified was a sequence between amino acid 411 to 490 of the herpes simplex VP16 protein. The amplified fragment was flanked with SpeI recognition sites. SpeI produces compatible ends to those of NheI sites. The oligonucleotides used :
VP16C 5' attactagtctgcggccccccgaccat 3' and
VP16E 5' aattactagtcccaccgtactcgtaattcc 3'
produced a 180 bp fragment which was restriction enzyme digested with SpeI and introduced 35 into pSKGRAHEcR NheI vector to produce pSKGRVP16HEcR. The DNA from the latter was sequenced and found to be mutation free, the junctions were also shown to be in frame with those of the glucocorticoid receptor.

The 2.2 kb EcoRV/NotI GRVP16HEcR fragment was introduced into a pcDNA3 EcoRV/NotI vector resulting in pcDNA3GRVP16HEcR (Figure 36).

Construction of plant transient expression effector plasmids containing the chimeric receptor with VP16 sequences

5 The same procedure was carried out to clone the GRVP16HeCR DNA fragment into tobacco(pMF7b) and maize(pMF6) expression vectors. A 2.2kb BamHI DNA fragment was isolated from pcDNA3GRVP16HeCR and ligated in to the pMF6 BamHI (or pMF7b BamHI) vector to produce pMF6GRVP16HeCR (Figure 37) (or pMF7GRVP16HeCR) (Figure 38).

Results - Addition of strong activation domains enhance ecdysone switch system

10 The VP16 transactivation domain from herpes simplex virus has been added to a maize protoplast vector pMF6GRHEcR to generate the vector pMF6GRVP16HEcR. When co-transformed into maize protoplasts with the reporter construct p221.9GRE6, in the presence of 100µM RH5992, enhanced levels of expression were seen over pMF6GRHEcR. Figure 39, shows that RH5992 is able to induce GUS levels comparable to those observed
15 with the positive control (p35SCaMVGUS), moreover, a dose response effect is observable.

VP16 enhanced vectors (pES3 and pES4) have been generated for stable transformation of tobacco. Following transformation transgenic progeny containing pES3 and pES4, gave elevated GUS levels following treatment with RH5992, relative to comparable transgenic plants containing the non-VP16 enhanced vectors pES1 and pES2.

20 An enhanced mammalian vector pcDNA3GRVP16HEcR was prepared for transient transfection of mammalian cell lines. Elevated reporter gene activities were obtained relative to the effector construct (pCDNA3GRHEcR) without the VP16 addition.

"Acidic" activation domains are apparently "universal" activators in eukaryotes (Ptashne (1988) Nature 335 683-689). Other suitable acidic activation domains which have
25 been used in fusions are the activator regions of GAL4 itself (region I and region II; Ma and Ptashne (Cell (1987) 48, 847-853), the yeast activator GCN4 (Hope and Struhl (1986) Cell 46, 885-894) and the herpes simplex virus VP16 protein (Triezenberg et al (1988) Genes Dev. 2, 718-729 and 730-742).

Other acidic and non-acidic transcriptional enhancer sequences for example from plant
30 fungal and mammalian species can be added to the chimeric ecdysone receptor to enhance induced levels of gene expression.

Chimeric or synthetic activation domains can be generated to enhance induced levels of gene expression.

Example X - Optimisation by replacement of *Heliothis* ligand binding domain in chimeric effector for that of an ecdysone ligand binding domain of another species

5 Mutagenesis of the ecdysone ligand binding domain results in the increased sensitivity
of the chimeric receptor for activating chemical. This can be achieved by deletions in the
ligand binding domain, use of error prone PCR (Caldwell et al., PCR Meth. Applic 2, 28-33
1992), and in vitro DNA shuffling PCR (Stemmer, Nature 370, 389-391 1994). To enhance
the efficacy of the listed techniques we have developed a screening system for enhanced levels
10 of induced expression (see below).

An alternative strategy to the mutation of a known ligand binding domain is to identify insect species which are particularly sensitive to ecdysteroid agonists. For example *Spodoptera exigua* is particularly sensitive to RH 5992. To investigate the role of the ecdysone receptor ligand binding domain in increased sensitivity to RH5992 we have isolated corresponding DNA sequences from of *S. exigua* (Figure 40, Sequence ID No. 6). Figure 41, Sequence ID No. 7 shows a protein alignment of the hinge and ligand binding domains of the *Heliothis virescens* and *Spodoptera exigua* ecdysone receptors. The protein sequence between the two species is well conserved.

Results - Manipulation of the ligand binding domain leads to enhanced induced expression

20 Isolation of an ecdysone ligand binding domain from another lepidopteran species was
carried out by using degenerate oligonucleotides and PCR of first strand cDNA (Perkin
Elmer, cDNA synthesis Kit) of the chosen species. The degenerate oligonucleotides at the 5'
end were HingxhoA and B and at the 3' end ligandxA/B

ligandxB 5' ttactcgagiacgtcccataatctcctciaaga 3'
 a t t a

- 1994). The first strand cDNA was used in PCR reactions under the following conditions
20mM Tris-HCL (pH8.4), 50mM KCl, 1.5mM MgCl₂, 200mM dNTPs (dATP,dCTP,dGTP
and dTTP) and 0.02 U/ml *Taq* DNA polymerase and in the presence of 1ug of each Hinge (5'
3') and Ligand (5'3') oligonucleotides. The PCR cycling conditions were 94°C for 1 minute,
5 60°C for 2 minutes and 72°C for 1 minute and 35 cycles were carried out. A sample of the
completed reaction was fractionated in a 1% agarose (w/v) 1 x TBE gel, and the resulting 900
bp fragment was subcloned into pCRII vector (Invitrogen). The resulting clone (pSKSEcR 1-
10) were further characterised and sequenced.
- 10 Example X - Manipulation of reporter gene promoter regions can modulate chemical
induced expression

The context of the effector response element in the reporter gene promoter can be
used to modulate the basal and induced levels of gene expression. Six copies of the
15 glucocorticoid response element were fused to 46 bp or 60 bp of the CaMV 35S promoter
sequence. When used with the effector construct pMF7GRHEcRS the reporter gene
construct containing 46 bp of the CaMV 35S promoter gave reduced basal and induced levels
of GUS expression relative to the 60 bp reporter construct (Figure 42).

20 Constructs for plant transformation (pES1 and ES2) have been generated to
demonstrate the size of minimal promoter can be used to modulate the basal and induced
levels of gene expression in plants.

The number and spacing of response elements in the reporter gene promoter can be
adjusted to enhance induced levels of trans-gene expression.

25 The utility of a two component system (effector and reporter gene cassettes) allows
the spatial control of induced expression. Trans-gene expression can be regulated in an tissue
specific, organ specific or developmentally controlled manner. This can be achieved by
driving the effector construct from a spatially or temporally regulated promoter.

References

- 30 Allan, G.F., Tsai, S.Y., Tsai, M.-J. and O'Malley, B.W. (1992a) P.N.A.S. **89**, 11750-11754.
Allan, G.F., Leng, X., Tsai, S.Y., Weigel, N.L., Edwards, D.P., Tsai, M.-J. and O'Malley,
B.W. (1992b) J. Biol. Chem **267**, 19513-19520.
Ashburner, M (1990) Cell **61**, 1-3.
35 Beato, M. (1989) Cell **56**, 335-344.
Carlberg, C., Bendik, I., Wyss, A., Meier, E., Sturzenbecker, L.J., Grippo, J.F. and Hunziker,
W. (1993) Nature **361**, 657-660.

- Christopherson, K.S., Mark., M.R., Bajaj, V. and Godowski, P.J. (1992) P.N.A.S. **89**, 6314-6318.
- Evans, R.M. (1988) Science **240**, 889-895.
- Green, S. and Chambon, P. (1988) TIGs **11**, 309-314.
- 5 Heyman, R.A., Mangelsdorf, D.J., Dyck, J.A., Stein, R.B., Eichele, G., Evans, R.M. and Thaller, C. (1992) Cell **68**, 397-406.
- Hirst, M.C., Bassett, J.H.D., Roche, A. and Davies, K.E. (1992) Trends in Genetics **8**, 6-7.
- Hogness, D.S., Talbot, W.S., Bender, M.T. and Koelle, M. (1992) X Ecdysone Workshop, Liverpool. Abstract.
- 10 Hollenberg, S.M., Weinberger, C., Ong, E.S., Cerelli, G., Oro, A., Lebo, R., Thompson, E.B., Rosenfeld, M.G. and Evans, R.M. (1985) Nature **318**, 635-641.
- Kliwera, S.A., Umesono, K., Mangeldorf, D.J. and Evans, R.M. (1992) Nature **355**, 446-449.
- Koelle, M.R., Talbot, W.S., Segraves, W.A., Bender, M.T., Cherbas, P. and Hogness, D.S.
- 15 (1991) Cell **67**, 59-77.
- Krust et al, (1986) The EMBO Journal **5**, 891-897.
- Leid, M., Kastner, P., Lyons, R., Nakshatri, H., Saunders, M., Zacharewski, T., Chen, J-Y., Staud, A., Garnier, J-M., Mader, S. and Chambon, P. (1992a) Cell **68**, 377-395.
- Leid, M., Kastner, P and Chambon, P. (1992b) TIBs **17**, 427-433.
- 20 Mangelsdorf, D.J., Borgmeyer, V., Heymann, R.A., Zhou, J.Y., Ong, E.S., Oro, A.E., Kakizuka, A. and Evans, R.M. (1992) Genes and Development **6**, 329-344.
- Oro, A.E., McKeown, M. and Evans, R.M. (1990) Nature **347**, 298-301.
- Riddihough, G. and Pelham, H.R.B. (1987) EMBO Journal **6**, 3729-3734.
- Segraves, W.A. (1991) Cell **67**, 225-228.
- 25 Segraves, W.A. and Hogness, D.S. (1990) Genes and Development **4**, 204-219.
- Smagghe, G. and Degheele, D (1994) Pestic. Sci. **42**, 85-92.
- Stemmer, W.P. (1994) Nature **370**, 389-391.
- Thummel, C.S., Burtis, K.S. and Hogness, D.S. (1990) Cell **61**, 101-111.
- Vegeta , E., Allan, G.F., Schrader, W. T., Tsai, M-J., McDonnell, D.P. and O'Maley, B.W.
- 30 (1992) Cell **69**, 703-713
- Yao, T.P., Segraves, W.A., Oro, A.E., McKeown, M. and Evans, R.M. (1992) Cell **71**, 63-72.
- Yao, T-P., Forman, B.M., Jlang, Z., Cherbas, L., Chen, J-Don., McKeown, M., Cherbas, P. and Evans, R.M. (1993) Nature **366**, 476-479.
- Yu, V.C., Delsert, C., Andersen, B., Holoway, J.M., Kim, S.Y., Boutin, J-M., Glass, C.K.
- 35 and Rosenfeld, M.G. (1991) Cell **67**, 1251-1266.

- 40 -

SEQUENCE LISTING

(1) GENERAL INFORMATION:

5

(i) APPLICANT:

- (A) NAME: ZENECA LIMITED
- (B) STREET: 15 STANHOPE GATE
- (C) CITY: LONDON
- 10 (E) COUNTRY: UK
- (F) POSTAL CODE (ZIP): W1Y 6LN

(ii) TITLE OF INVENTION: A GENE SWITCH

15

(iii) NUMBER OF SEQUENCES: 7

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- 20 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(vi) PRIOR APPLICATION DATA:

25

- (A) APPLICATION NUMBER: GB 9510759.5
- (B) FILING DATE: 26-MAY-1995

(vi) PRIOR APPLICATION DATA:

30

- (A) APPLICATION NUMBER: GB 9513882.3
- (B) FILING DATE: 07-JUL-1995

(vi) PRIOR APPLICATION DATA:

35

- (A) APPLICATION NUMBER: GB 9517316.7
- (B) FILING DATE: 24-AUG-1995

(vi) PRIOR APPLICATION DATA:

40

- (A) APPLICATION NUMBER: GB 9605656.9
- (B) FILING DATE: 18-MAR-1996

40 (2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

45

- (A) LENGTH: 116 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

50

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Heliothis virescens

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

TGCAGGGGT GCAAGGAGTT CTTCAGGCCG AGTGTAACCA AAAATGCAGT GTACATATGC

60

AAATTCGGCC ATGCTTGCGA AATGGATATG TATATGCGGA GAAAATGCCA AGAGTA

116

60

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- 5
- (A) LENGTH: 1934 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: circular

10

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Heliothis virescens

15

(vii) IMMEDIATE SOURCE:
(B) CLONE: pSK19R

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

TCCACTGGTG	TTTCACCAC	CACAGAAAAG	GCCTCTGCTC	ATTTAGAGGG	TGGTGCTAAG	60	
20	AAGGTCACTCA	TCTCCTGCTG	CCCACCGCTG	ACCCATGTTC	GTCGTTGGTG	TCAACCTTGA	120
	AGCAGTATGA	CCCCCTCTTAC	AAGGTCACTCT	CCAACGCCTC	CTGCACAAACC	AACTGCCTCG	180
25	CTCCTCTCGC	TAAGGTCACTC	CATGACAAC	TCGAGATCAT	TGAAGGTCTG	ATGACCACTG	240
	TACACGCCAC	CACTGCCACC	CAGAACAGACAG	TGGATGGACC	CTCTGGTAAA	CTGTGGCGTG	300
	ATGGCCGTGG	TGCTCAGCAG	AATATCATTTC	CCGCGGAATT	CCCCAGCCGC	AGCTAGCTAA	360
30	CCTGCAGCAG	ACACAAACCCC	TACCTTCCAT	GCCGTTACCA	ATGCCACCGA	CAACACCCAA	420
	ATCAGAAAAC	GAGTCAATGT	CATCAGGTCTG	TGAGGAACTG	TCTCCAGCTT	CGAGTGTAAA	480
35	CGGCTGCAGC	ACAGATGGCG	AGGCGAGGCG	GCAGAAGAAA	GGCCCAGCGC	CGAGGCAGCA	540
	AGAAAGAGCTA	TGTCTTGTCT	CGGGCGACAG	AGCCTCCGGA	TATCACTACA	ACCGCGCTCAC	600
	ATGTGAAGGG	TGTAAAGGTT	TCTTCAGGCG	GAGTGTAAACC	AAAAATGCAG	TGTACATATG	660
40	CAAATT CGGC	CATGCTT GCG	AAATGGATAT	CTATATGCGG	AGAAAATGTC	AGGAGTGTCTG	720
	GTTGAAGAAA	TGTCTT GCGG	TGGGCATGAG	GCCCGAGTGC	GTGGTGCCGG	AGAACCCAGTG	780
45	TGCAATGAAA	CGGAAAGAGA	AAAAGGCAGCA	GAGGGAAAAAA	GACAAATTGC	CCGTCAGTAC	840
	GACGACAGTA	GACCGATCACA	TGCTCTCCAT	CATGCAATGT	GACCCCTCCGC	CCCCAGAGGC	900
	CGCTAGAATT	CTGGAATGTG	TGCAGCACGA	GGTGGTGCCCA	CGATT CCTGA	ATGAGAAGCT	960
50	AATGGAACAG	AACAGATTGA	AGAACGTGCC	CCCCCTCACT	GCCAATCAGA	AGTCGTTGAT	1020
	CGCAAGGCTC	GTGTGGTACC	AGGAAGGCTA	TGAACAAACCT	TCCGAGGAAG	ACCTGAAGAG	1080
55	GGTTACACAG	TCGGACGAGG	ACGACGAAGA	CTCGGATATG	CCGTTCCGTC	AGATTACCGA	1140
	GATGACGATT	CTCACAGTGC	AGCTCATCGT	AGAATTGCT	AAGGGCCTCC	CGGGCTTCGC	1200
	CAAGATCTCG	CAGTCGGACC	AGATCACGTT	ATTAAGGCG	TGCTCAAGTG	AGGTGATGAT	1260
60	GCTCCGAGTG	GCTCGGCCGT	ATGACGCGGC	CACCGACAGC	GTACTGTTCG	CGAACAAACCA	1320
	GGCGTACACT	CGCGACAACT	ACCGCAAGGC	AGGCATGGCG	TACGTCACTG	AGGACCTGCT	1380

	GCACTTCTGT CGGTGCATGT ACTCCATGAT GATGGATAAC GTGCATTATG CGCTGCTTAC	1440
	AGCCATTGTC ATCTTCTCAG ACCGGCCCCG GCTTGAGCAA CCCCTGTTGG TGGAGGACAT	1500
5	CCAGAGATAT TACCTGAACA CGCTACGGGT GTACATCCTG AACCAGAACCA GCGCGTCGCC	1560
	CCGGGGCGCC GTCATCTCG GCGAGATCCT GGGCATACTG ACGGAGATCC GCACGCTGGG	1620
10	CATGCAGAAC TCCAACATGT GCATCTCCCT CAAGCTGAAG AACAGGAAGC TGCCGCCGTT	1680
	CCTCGAGGAG ATCTGGGACG TGGCGGACGT GGCGACGACG GCGACGCCGG TGGCGGCCGGA	1740
	GGCGCCGGCG CCTCTAGCCC CCGCCCCGCC CGCCCGGCCG CCCGCCACCG TCTAGCGCGC	1800
15	CTCAGGAGAG AACGCTCATA GACTGGCTAG TTTTAGTGAA GTGCACGGAC ACTGACGTCG	1860
	ACGTGATCAA CCTATTATA AGGACTGCGA ATTTTACAC TTAAGAGGGC ACACCCGTAC	1920
20	CCGATTTCGT ACGG	1934

(2) INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2464 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: cDNA
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: *Heliothis virescens*
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: pSK16.1

40	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:	
	CGCTGGTATA ACAACGGACC ATTCCAGACG CTGCGAATGC TCGAGGAGAG CTCGTCTGAG	60
	GTGACGTCGT CTTCAGCACT GGGCCTGCCG CCGGCTATGG TGATGTCCCC GGAATCGCTC	120
45	GCGTCGCCCG AGATCGGCCG CCTGGAGCTG TGGGGCTACG ACGATGGCAT CACTTACAGC	180
	ATGGCACAGT CGCTGGGCAC CTGCACCATG GAGCAGCAGC AGCCCCAGCC GCAGCAGCAG	240
50	CCGCAGCAGA CACAACCCCT ACCTTCCATG CCGTTACCAA TGCCACCGAC AACACCCAAA	300
	TCAGAAAACG AGTCAATGTC ATCAGGTCGT GAGGAACGTG CTCCAGCTTC GAGTGAAAC	360
	GGCTGCAGCA CAGATGGCGA GGCGAGGCCG CAGAAGAAAG GCCCAGCGCC GAGGCAGCAA	420
55	GAAGAGCTAT GTCTTGTCTG CGGGCAGACA GCCTCCGGAT ATCACTACAA CGCGCTCACA	480
	TGTGAAGGGT GTAAAGGTTT CTTCAGGCCG AGTGTAAACCA AAAATGCACT GTACATATGC	540
60	AAATTCCGGC ATGCTTGCAG AATGGATATC TATATGCCGA GAAAATGTCA GGAGTGTCCG	600
	TTGAAGAAAT GTCTTGCAGT GGGCATGAGG CCCGAGTGCAG TGGTGCCCGA GAACCAAGTGT	660
	GCAATGAAAC GGAAAGAGAA AAAGGCGCAG AGGGAAAAAG ACAAAATTGCC CGTCAGTACG	720

	ACGACAGTAG ACGATCACAT GCCTCCCAC	ATGCAATGTG ACCCTCCGCC CCCAGAGGCC	780
5	GCTAGAACATTG	TGGAATGTGT GCAGCACGAG GTGGTGCCAC GATTCTGAA TGAGAAGCTA	840
	ATGGAACAGA ACAGATTGAA GAACGTGCC	CCCCTCACTG CCAATCAGAA GTCGTTGATC	900
	GCAAGGCTCG TGTGGTACCA GGAAGGCTAT	GAACAACCTT CCGAGGAAGA CCTGAAGAGG	960
10	GTTACACAGT CGGACGAGGA CGACGAAGAC	TCGGATATGC CGTTCCGTCA GATTACCGAG	1020
	ATGACGATTG TCACAGTGCA	GCTCATCGTA GAATTGCTA AGGGCCTCCC GGGCTTCGCC	1080
15	AAGATCTCGC AGTCGGACCA GATCACGTTA	TTAAAGGCGT GCTCAAGTGA GGTGATGATG	1140
	CTCCGAGTGG CTGGCGGTA TGACGGGCC	ACCGACAGCG TACTGTTCGC GAACAACCAG	1200
	CGGTACACTC GCGACAACTA CCGCAAGGCA	GGCATGGCGT ACGTCATCGA GGACCTGCTG	1260
20	CACTTCTGTC GGTGCATGTA	CTCCATGATG ATGGATAACG TGCATTATGC GCTGCTTACA	1320
	GCCATTGTCA TCTTCTCAGA CCGGCCCCGG	CTTGAGCAAC CCCTGTTGGT GGAGGACATC	1380
	CAGAGATATT ACCTGAACAC GCTACGGGTG	TACATCCTGA ACCAGAACAG CGCGTCGCC	1440
25	CGCGGCGCCG TCATCTTCGG	CGAGATCCTG GGCATACTGA CGGAGATCCG CACGCTGGC	1500
	ATGCAGAACT CCAACATGTG CATCTCCCTC	AAGCTGAAGA ACAGGAAGCT GCCGCCGTT	1560
30	CTCGAGGAGA TCTGGGACGT	GGCGGACGTG GCGACGACGG CGACGCCGGT GGCGGCGGAG	1620
	GCGCCGGCGC CTCTAGCCCC CGCCCCGCC	GCCCCGCCGC CGGCCACCGT CTAGCGCGCC	1680
35	TCAGGAGAGA ACGCTCATAG ACTGGCTAGT	TTTAGTGAAG TGACGGACA CTGACGTCGA	1740
	CGTGATCAAC CTATTTATAA GGACTGCGAA	TTTTACCACT TAAGAGGGCA CACCCGTACC	1800
	CGATTTCGTA CGTATTGGGT GACCGACGAC	GATGCAGAGC GTGTGTAATG TGAATATATG	1860
40	TGTTGTTGAA CGATTTGGAG AATATATATT	GGTGGTGTG TTGGGCCCG CACGCCGTCG	1920
	CCGGTGGCG GCGATCGCGG CGCCCCGCC	TTCAGTTTA TTTCGTTTAC GACTGAGTTG	1980
	GTCACTCGGA TACGACTGTA	TGATAAGACT TCGTTCGATA AGTACACCTA CTAATTAC	2040
45	CATACGTACG TAGCTTACGA GAGTTATTAG	AGACAAAGAA TATAAGAAGA AGATGTTCT	2100
	ATTGGGTGAA AAGTTGATAG TTATGTTAT	TTACAAAAAT TAACAATAAT ACGTTGATTA	2160
50	ACCTTCGAG TATAATATTG TGATGAGTCG	TCCGCTGTCC ACGTCGCCGT CACATGTTG	2220
	TTTCTGATGC ACACGTGAGG NGCGTTATCG	TGTTTCATGG TTCCATCGTC CTGTGCC	2280
55	GACCCCTCGAC TAAATGAGTA ATTAAATTAA	TTGCTGTGAT TACATTTAA TGTGTTGATT	2340
	ATCTACCATA GGGTGATATA AGTGTGTCTT	ATTACAATAA AAAGTGTGTG TCGTCGATAG	2400
	CTTCCACACG AGCAAGCCTT TTGTTAAAGT	GATTTACTGA CATGGACACT CGACCCGGAA	2460
60	CTTC		2464

(2) INFORMATION FOR SEQ ID NO: 4:

5

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2745 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: cDNA

15 (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 225..1955
- (D) OTHER INFORMATION: /codon_start= 225
/product= "Heliothis ecdysone receptor"

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

20	ACTCGCGTGC TCTTCTCAC C TGTTGCTCGG ATTGTGTTGT ACTAGAAAAAA AGTTGTGCC	60
	GCTCGAACGA GACTTCCGAG TCCTATTGGA TTGCACGAAA GTCGAGACAG TGGATAGCGA	120
25	TTCGGTTTCG TTTGAACCTT GCGTAGACGA GTGGTGCATG TCCATGAGTC GCGTTTAGAT	180
	AGTTTAGTGC GAGGAAAAAG TGAAGTGAAA GCCTTCCTCG GAGGATGTCC CTCGGCGCTC	240
	GTGGATAACCG GAGGTGTGAC ACGCTCGCCG ACATGAGACG CCGCTGGTAT AACAACGGAC	300
30	CATTCCAGAC GCTGCGAATG CTCGAGGAGA GCTCGTCTGA GGTGACGTCG TCTTCAGCAC	360
	TGGGCCTGCC GCCGGCTATG GTGATGTCCC CGGAATCGCT CGCGTCGCCG GAGATCGGCG	420
35	GCCTGGAGCT GTGGGGCTAC GACGATGGCA TCACTTACAG CATGGCACAG TCGCTGGCA	480
	CCTGCACCAT GGAGCAGCAG CAGCCCCAGC CGCAGCAGCA GCCGCAGCAG ACACAACCCC	540
	TACCTTCCAT GCCGTTACCA ATGCCACCGA CAACACCCAA ATCAGAAAAC GAGTCAATGT	600
40	CATCAGGTCG TGAGGAAC TG TCTCCAGCTT CGAGTGTAAGA CGGCTGCAGC ACAGATGGCG	660
	AGGCGAGGCG GCAGAAGAAA GGCCAGCGC CGAGGCAGCA AGAAGAGCTA TGTCTTGTCT	720
45	CGCGCGACAG AGCCTCCGGA TATCACTACA ACGCGCTCAC ATGTGAAGGG TGTAAAGGTT	780
	TCTTCAGGCG GAGTGTAACC AAAATGCAG TGTACATATG CAAATTCCGC CATGCTTGC	840
	AAATGGATAT CTATATGCCG AGAAAATGTC AGGAGTGTCC GTTGAAGAAA TGTCTTGC	900
50	50 TGGGCATGAG GCCCGAGTGC GTGGTCCGG AGAACCAAGTG TGCAATGAAA CGGAAAGAGA	960
	AAAAGGCGCA GAGGGAAAAA GACAAATTGC CCGTCAGTAC GACGACAGTA GACGATCACA	1020
55	TGCCCTCCAT CATGCAATGT GACCCCTCCGC CCCCAGAGGC CGCTAGAATT CTGGAATGTG	1080
	TGCAGCACGA GGTGGTGCCTA CGATTCCGTA ATGAGAAGCT AATGGAACAG AACAGATTGA	1140
	AGAACGTGCC CCCCTCACT GCCAATCAGA AGTCGTTGAT CGCAAGGCTC GTGTGGTACC	1200
60	AGGAAGGCTA TGAACAAACCT TCCGAGGAAG ACCTGAAGAG GGTTACACAG TCGGACGAGG	1260
	ACGACGAAGA CTCGGATATG CCGTTCCGTC AGATTACCGA GATGACGATT CTCACAGTGC	1320

- 45 -

	AGCTCATCGT AGAATTGCGT AAGGGCCTCC CGGGCTTCGC CAAGATCTCG CAGTCGGACC	1380
	AGATCACGTT ATTAAAGGCG TGCTCAAGTG AGGTGATGAT GCTCCGAGTG GCTCGGGCGT	1440
5	ATGACGCGGC CACCGACAGC GTACTGTTCG CGAACAAACCA GGCGTACACT CGCGACAACT	1500
	ACCGCAAGGC AGGCATGGCG TACGTACATCG AGGACCTGCT GCACCTCTGT CGGTGCATGT	1560
10	ACTCCATGAT GATGGATAAC GTGCATTATG CGCTGCTTAC AGCCATTGTC ATCTTCTCAG	1620
	ACCGGCCCGG GCTTGAGCAA CCCCTGTTGG TGGAGGAGAT CCAGAGATAT TACCTGAACA	1680
	CGCTACGGGT GTACATCCG AACAGAACAGA GCGCGTCGCC CGCGCGGCC GTCATCTTCG	1740
15	GCGAGATCCT GGGCATACTG ACGGAGATCC GCACGCTGGG CATGCAGAAC TCCAACATGT	1800
	GCATCTCCCT CAAGCTGAAG AACAGGAAGC TGCCGCCGTT CCTCGAGGAG ATCTGGGACG	1860
20	TGGCGGACGT GGCGACGACG GCGACGCCGG TGGCGGCCGG GGCGCCGGCG CCTCTAGCCC	1920
	CCGCCCCGCC CGCCCGGCCG CCCGCCACCG TCTAGCGCGC CTCAGGAGAG AACGCTCATA	1980
	GACTGGCTAG TTTTAGTGAA GTGCACGGAC ACTGACGTCG ACGTGATCAA CCTATTTATA	2040
25	AGGACTGCGA ATTTTACAC TTAAGAGGGC ACACCCGTAC CCGATTTCGT ACGTATTCGG	2100
	TGACCGACGA CGATGCAGAG CGTGTGTAAT GTGAATATAT GTGTTGTTGA ACGATTTGGA	2160
	GAATATATAT TGGTGTGCT GTTCGGGCC GCACGCCGTC GCCGGTCGGC GGCGATCGCG	2220
30	GCGCCCGCGG CTTCAGTTT ATTTGTTTA CGACTGAGTT GGTCACTCGG ATACGACTGT	2280
	ATGATAAGAC TTGTTCCAT AAGTACACCT ACTAAATTAC ACATACGTAC GTAGCTTACG	2340
35	AGAGTTATTA GAGACAAAGA ATATAAGAAG AAGATGTTTC TATTGGGTGA AAAGTTGATA	2400
	GTTATGTTTA TTTACAAAAA TTAACAATAA TACGTTGATT AACCTTCGA GTATAATATT	2460
40	GTGATGAGTC GTCCGCTGTC CACGTCGCCG TCACATGTTT GTTCTGATG CACACGTGAG	2520
	GNCGCTTATC GTGTTTCATG GTTCCATCGT CCTGTGCCCG CGACCCCTCGA CTAATGAGT	2580
	AATTTAATTG ATTGCTGTGA TTACATTTA ATGTGTTGAT TATCTACCAT AGGGTGATAT	2640
45	AAGTGTGTCT TATTACAATA CAAAGTGTGT GTCGTCGATA GCTTCCACAC GAGCAAGCCT	2700
	TTTGTAAAG TGATTTACTG ACATGGACAC TCGACCCGGG ACTTC	2745

50 (2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 575 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- 55 (ii) MOLECULE TYPE: protein

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

- 46 -

	Met Ser Leu Gly Ala Arg Gly Tyr Arg Arg Cys Asp Thr Leu Ala Asp	
	1 5 10 15	
5	Met Arg Arg Arg Trp Tyr Asn Asn Gly Gly Phe Gln Thr Leu Arg Met	
	20 25 30	
	Leu Glu Glu Ser Ser Ser Glu Val Thr Ser Ser Ser Ala Leu Gly Leu	
	35 40 45	
10	Pro Pro Ala Met Val Met Ser Pro Glu Ser Leu Ala Ser Pro Glu Ile	
	50 55 60	
	Gly Gly Leu Glu Leu Trp Gly Tyr Asp Asp Gly Ile Thr Tyr Ser Met	
	65 70 75 80	
15	Ala Gln Ser Leu Gly Thr Cys Thr Met Glu Gln Gln Gln Pro Gln Pro	
	85 90 95	
20	Gln Gln Gln Pro Gln Gln Thr Gln Pro Leu Pro Ser Met Pro Leu Pro	
	100 105 110	
	Met Pro Pro Thr Thr Pro Lys Ser Glu Asn Glu Ser Met Ser Ser Gly	
	115 120 125	
25	Arg Glu Glu Leu Ser Pro Ala Ser Ser Val Asn Gly Cys Ser Thr Asp	
	130 135 140	
	Gly Glu Ala Arg Arg Gln Lys Lys Gly Pro Ala Pro Arg Gln Gln Glu	
	145 150 155 160	
30	Glu Leu Cys Leu Val Cys Gly Asp Arg Ala Ser Gly Tyr His Tyr Asn	
	165 170 175	
35	Ala Leu Thr Cys Glu Gly Cys Lys Gly Phe Phe Arg Arg Ser Val Thr	
	180 185 190	
	Lys Asn Ala Val Tyr Ile Cys Lys Phe Gly His Ala Cys Glu Met Asp	
	195 200 205	
40	Ile Tyr Met Arg Arg Lys Cys Gln Glu Cys Arg Leu Lys Lys Cys Leu	
	210 215 220	
	Ala Val Gly Met Arg Pro Glu Cys Val Val Pro Glu Asn Gln Cys Ala	
	225 230 235 240	
45	Met Lys Arg Lys Glu Lys Lys Ala Gln Arg Glu Lys Asp Lys Leu Pro	
	245 250 255	
50	Val Ser Thr Thr Val Asp Asp His Met Pro Pro Ile Met Gln Cys	
	260 265 270	
	Asp Pro Pro Pro Glu Ala Ala Arg Ile Leu Glu Cys Val Gln His	
	275 280 285	
55	Glu Val Val Pro Arg Phe Leu Asn Glu Lys Leu Met Glu Gln Asn Arg	
	290 295 300	
	Leu Lys Asn Val Pro Pro Leu Thr Ala Asn Gln Lys Ser Leu Ile Ala	
	305 310 315 320	
60	Arg Leu Val Trp Tyr Gln Glu Gly Tyr Glu Gln Pro Ser Glu Glu Asp	
	325 330 335	

- 47 -

Leu Lys Arg Val Thr Gln Ser Asp Glu Asp Asp Glu Asp Ser Asp Met
 340 345 350

5 Pro Phe Arg Gln Ile Thr Glu Met Thr Ile Leu Thr Val Gln Leu Ile
 355 360 365

Val Glu Phe Ala Lys Gly Leu Pro Gly Phe Ala Lys Ile Ser Gln Ser
 370 375 380

10 Asp Gln Ile Thr Leu Leu Lys Ala Cys Ser Ser Glu Val Met Met Leu
 385 390 395 400

Arg Val Ala Arg Arg Tyr Asp Ala Ala Thr Asp Ser Val Leu Phe Ala
 405 410 415

15 Asn Asn Gln Ala Tyr Thr Arg Asp Asn Tyr Arg Lys Ala Gly Met Ala
 420 425 430

20 Tyr Val Ile Glu Asp Leu Leu His Phe Cys Arg Cys Met Tyr Ser Met
 435 440 445

Met Met Asp Asn Val His Tyr Ala Leu Leu Thr Ala Ile Val Ile Phe
 450 455 460

25 Ser Asp Arg Pro Gly Leu Glu Gln Pro Leu Leu Val Glu Asp Ile Gln
 465 470 475 480

Arg Tyr Tyr Leu Asn Thr Leu Arg Val Tyr Ile Leu Asn Gln Asn Ser
 485 490 495

30 Ala Ser Pro Arg Gly Ala Val Ile Phe Gly Glu Ile Leu Gly Ile Leu
 500 505 510

35 Thr Glu Ile Arg Thr Leu Gly Met Gln Asn Ser Asn Met Cys Ile Ser
 515 520 525

Leu Lys Leu Lys Lys Arg Lys Leu Pro Pro Phe Leu Glu Glu Ile Trp
 530 535 540

40 Asp Val Ala Asp Val Ala Thr Thr Ala Thr Pro Val Ala Ala Glu Ala
 545 550 555 560

Pro Ala Pro Leu Ala Pro Ala Pro Pro Ala Arg Pro Ala Thr Val
 565 570 575

45 (2) INFORMATION FOR SEQ ID NO: 6:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 948 base pairs
 50 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

55 (ii) MOLECULE TYPE: cDNA
 (vi) ORIGINAL SOURCE:
 (A) ORGANISM: Spodoptera exigua

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
 AGGCCGGAGT GCGTGGTGCC AGAAAACCAG TGTGCAATGA AAAGGAAAGA GAAAAAGGCA

	CAAAGGGAAA AAGACAAGTT GCCAGTCAGT ACAACGACAG TGGATGATCA CATGCCTCCC	120
5	ATTATGCAGT GTGATCCACC GCCTCCAGAG GCCGCAAGAA TTCACGAGGT GGTGCCACGA	180
	TTCCTGAATG AAAAGCTAAT GGACAGGACA AGGCTCAAGA ATGTGCCCCC TCACTGCCAA	240
	CCAGAAAGTCC TTAATAGCGA GGCTGGTCTG GTACCAAGAA GGCTATGAAC AGCCATCAGA	300
10	AGAGGATCTA AAAAGAGTCA CACAGTCGGA TGAAGACGAA GAAGAGTCGG ACATGCCGTT	360
	CCGTCAGATC ACCGAGATGA CGATCCTCAC AGTGCAGCTC ATTGTTGAAT TCGCTAAGGG	420
	CCTACCAGCG TTGGCAAAGA TCTCACAGTC GGATCAGATC ACATTATTAA AGGCCTGTTTC	480
15	GAGTGAGGTG ATGATGTTGC GAGTAGCTCG GCGGTACGAC GCGGCGACAG ACAGCGTGTT	540
	GTTCGCCAAC AACCAAGGCGT ACACCCGCGA CAACTACCGC AAGGCAGGCA TGGCCTACGT	600
20	CATCGAGGAC CTGCTGCACT TCTGCCGGTG CATGTACTCC ATGATGATGG ATAACGTCCA	660
	CTATGCACTG CTCACTGCCA TCGTCATTTT CTCAGACCGA CCCGGGCTTG AGCTAACCCCT	720
	GTTGGTGGAG GAGATCCAGA GATATTACCT GAACACGCTG CGGGTGTACA TCCTGAACCA	780
25	GAACAGTCGG TCGCCGTGCT GCCCTGTCACT CTACGCTAAG ATCCTCGGCA TCCTGACGGAA	840
	GCTGCGGACC CTGGGCATGC AGAACTCCAA CATGTGCATC TCACTCAAGC TGAAGAACAG	900
30	GAACGTGCGC CGGTTCTTCG AGGATATCTG GGACGTCCTC GAGTAAAA	940

(2) INFORMATION FOR SEO ID NO: 7:

- 35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 319 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

40 (ii) MOLECULE TYPE: protein

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Arg Pro Glu Cys Val Val Pro Glu Asn Gln Cys Ala Met Lys Arg Lys
1 5 10 15

50 Glu Lys Lys Ala Gln Arg Glu Lys Asp Lys Leu Pro Val Ser Thr Thr
20 25 30

Thr Val Asp Asp His Met Pro Pro Ile Met Gln Cys Asp Pro Pro Pro
35 40 45

Pro Glu Ala Ala Arg Ile Leu Glu Cys Val Gln His Glu Val Val Pro
50 55 60

60 Arg Phe Leu Asn Glu Lys Leu Met Glu Gln Asn Arg Leu Lys Asn Val
65 70 75 80

Pro Pro Leu Thr Ala Asn Gln Lys Ser Leu Ile Ala Arg Leu Val Trp
85 90 95

- 49 -

Tyr Gln Glu Gly Tyr Glu Gln Pro Ser Glu Glu Asp Leu Lys Arg Val
100 105 110

5 Thr Gln Ser Asp Glu Asp Asp Glu Asp Ser Asp Met Pro Phe Arg Gln
115 120 125

Ile Thr Glu Met Thr Ile Leu Thr Val Gln Leu Ile Val Glu Phe Ala
130 135 140

10 Lys Gly Leu Pro Gly Phe Ala Lys Ile Ser Gln Ser Asp Gln Ile Thr
145 150 155 160

Leu Leu Lys Ala Cys Ser Ser Glu Val Met Met Leu Arg Val Ala Arg
15 165 170 175

Arg Tyr Asp Ala Ala Thr Asp Ser Val Leu Phe Ala Asn Asn Gln Ala
180 185 190

20 Tyr Thr Arg Asp Asn Tyr Arg Lys Ala Gly Met Ala Tyr Val Ile Glu
195 200 205

Asp Leu Leu His Phe Cys Arg Cys Met Tyr Ser Met Met Met Asp Asn
210 215 220

25 Val His Tyr Ala Leu Leu Thr Ala Ile Val Ile Phe Ser Asp Arg Pro
225 230 235 240

Gly Leu Glu Gln Pro Leu Leu Val Glu Glu Ile Gln Arg Tyr Tyr Leu
30 245 250 255

Asn Thr Leu Arg Val Tyr Ile Leu Asn Gln Asn Ser Ala Ser Pro Arg
260 265 270

35 Gly Ala Val Ile Phe Gly Glu Ile Leu Gly Ile Leu Thr Glu Ile Arg
275 280 285

Thr Leu Gly Met Gln Asn Ser Asn Met Cys Ile Ser Leu Lys Leu Lys
40 290 295 300

Lys Arg Lys Leu Pro Pro Phe Leu Glu Glu Ile Asp Trp Asp Val
305 310 315

CLAIMS

1. DNA comprising the sequence shown in Seq ID No. 2.
- 5 2. DNA comprising the sequence shown in Seq ID No. 3.
3. DNA comprising the sequence shown in Seq ID No. 4.
4. DNA comprising a sequence which shows 60% or more homology with the sequence
10 shown in Seq ID No 1, 2 or 3.
5. DNA according to claim 4 wherein said homology is in the range of 65% to 99%.
6. DNA which hybridises to the sequence shown in Seq. ID No. 2, 3 or 4, and which
15 codes for at least part of the *Heliothis* ecdysone receptor.....
7. DNA which is degenerate as a result of the genetic code to the DNA of any one of
claims 1 to 6 and which codes for a polypeptide which is at least part of the *Heliothis*
ecdysone receptor.
- 20 8. DNA comprising part of the sequence shown in Seq ID No. 2, and which codes for at
least part of the *Heliothis* ecdysone receptor ligand binding domain.
9. DNA comprising part of the sequence shown in Seq ID No. 3, and which codes for at
25 least part of the *Heliothis* ecdysone receptor ligand binding domain.
10. DNA comprising part of the sequence shown in Seq ID No. 4, and which codes for at
least part of the *Heliothis* ecdysone receptor ligand binding domain.
- 30 11. DNA comprising a sequence which shows 60% or more homology with the sequence
of claim 8, 9 or 10.
12. DNA according to claim 11 wherein said homology is in the range of 65% to 99%.
- 35 13. DNA which hybridises to the DNA of any one of claims 8 to 12 and which codes for
at least part of the *Heliothis* ecdysone receptor ligand binding domain.

14. DNA which is degenerate as a result of the genetic code to the DNA of any one of claims 8 to 12 and which codes for a polypeptide which is at least part of the *Heliothis* ecdysone receptor ligand binding domain.
- 5 15. DNA comprising part of the sequence shown in Seq ID No. 2, and which codes for at least part of the *Heliothis* ecdysone receptor DNA binding domain.
16. DNA comprising part of the sequence shown in Seq ID No. 3, and which codes for at least part of the *Heliothis* ecdysone receptor DNA binding domain.
- 10 17. DNA comprising part of the sequence shown in Seq ID No. 4, and which codes for at least part of the *Heliothis* ecdysone receptor DNA binding domain.
- 15 18. DNA comprising a sequence which shows 60% or more homology with the sequence of claim 15, 16 or 17.
19. DNA according to claim 18 wherein said homology is in the range of 65% to 99%.
- 20 20. DNA which hybridises to the DNA of any one of claims 15 to 19 and which codes for at least part of the *Heliothis* ecdysone receptor DNA binding domain.
21. DNA which is degenerate as a result of the genetic code to the DNA of any one of claims 15 to 19 and which codes for a polypeptide which is at least part of the *Heliothis* ecdysone receptor DNA binding domain.
- 25 22. DNA comprising part of the sequence shown in Seq ID No. 2, and which codes for at least part of the *Heliothis* ecdysone receptor transactivation domain.
23. DNA comprising part of the sequence shown in Seq ID No. 3, and which codes for at least part of the *Heliothis* ecdysone receptor transactivation domain.
- 30 24. DNA comprising part of the sequence shown in Seq ID No. 4, and which codes for at least part of the *Heliothis* ecdysone receptor transactivation domain.
- 35 25. DNA comprising a sequence which shows 60% or more homology with the sequence of claim 22, 23 or 24.

26. DNA according to claim 25 wherein said homology is in the range of 65% to 99%.
27. DNA which hybridises to the DNA of any one of claims 22 to 26 and which codes for at least part of the *Heliothis* ecdysone receptor transactivation domain.
5
28. DNA which is degenerate as a result of the genetic code to the DNA of any one of claims 22 to 26 and which codes for a polypeptide which is at least part of the *Heliothis* ecdysone receptor transactivation domain.
- 10 29. DNA comprising part of the sequence shown in Seq ID No. 2, and which codes for at least part of the *Heliothis* ecdysone receptor hinge domain.
30. DNA comprising part of the sequence shown in Seq ID No. 3, and which codes for at least part of the *Heliothis* ecdysone receptor hinge domain.
15
31. DNA comprising part of the sequence shown in Seq ID No. 4, and which codes for at least part of the *Heliothis* ecdysone receptor hinge domain.
- 20 32. DNA comprising a sequence which shows 60% or more homology with the sequence of claim 29, 30 or 31.
33. DNA according to claim 32 wherein said homology is in the range of 65% to 99%.
- 25 34. DNA which hybridises to the DNA of any one of claims 29 to 33 and which codes for at least part of the *Heliothis* ecdysone receptor hinge domain.
35. DNA which is degenerate as a result of the genetic code of the DNA of any one of claims 29 to 33 and which codes for a polypeptide which is at least part of the *Heliothis* ecdysone receptor hinge domain.
30
36. DNA having part of the sequence shown in Seq ID No. 2, and which codes for at least part of the *Heliothis* ecdysone receptor carboxy terminal region.
37. DNA having part of the sequence shown in Seq ID No. 3, and which codes for at least part of the *Heliothis* ecdysone receptor carboxy terminal region.
35

38. DNA having part of the sequence shown in Seq ID No. 4, and which codes for at least part of the *Heliothis* ecdysone receptor carboxy terminal region.
39. DNA comprising a sequence which shows 60% or more homology with the sequence of claim 36, 37 or 38.
40. DNA according to claim 39 wherein said homology is in the range of 65% to 99%.
41. DNA which hybridises to the DNA of any one of claims 36 to 40 and which codes for at least part of the *Heliothis* ecdysone receptor carboxy terminal region.
42. DNA which is degenerate as a result of the genetic code of the DNA of any one of claims 36 to 40 and which codes for a polypeptide which is at least part of the *Heliothis* ecdysone receptor carboxy terminal region.
43. A polypeptide comprising the *Heliothis* ecdysone receptor or a fragment thereof, wherein said polypeptide is substantially free from other proteins with which it is ordinarily associated, and which is coded for by the DNA of any preceding claim.
44. A polypeptide comprising the amino acid sequence shown in Seq ID No. 4 or any allelic variant or derivative thereof.
45. A polypeptide comprising part of the amino acid sequence shown in Seq ID No. 4 or any allelic variant or derivative thereof, which sequence provides the *Heliothis* ecdysone receptor ligand binding domain.
46. A polypeptide comprising part of the amino acid sequence shown in Seq ID No. 4 or any allelic variant or derivative thereof, which sequence provides the *Heliothis* ecdysone receptor DNA binding domain.
47. A polypeptide comprising part of the amino acid sequence shown in Seq ID No. 4 or any allelic variant or derivative thereof, which sequence provides the *Heliothis* ecdysone receptor transactivation domain.
48. A polypeptide comprising part of the amino acid sequence shown in Seq ID No. 4 or any allelic variant or derivative thereof, which sequence provides the *Heliothis* ecdysone receptor hinge domain.

49. A polypeptide comprising part of the amino acid sequence shown in Seq ID No. 4 or any allelic variant or derivative thereof, which sequence provides the *Heliothis* ecdysone receptor carboxy terminal region.
5
50. A polypeptide according to any one of claims 44 to 49 wherein said derivative is a homologous variant which includes conservative amino acid changes.
51. DNA comprising the sequence shown in Seq ID No. 6.
10
52. DNA comprising a sequence which shows 60% or more homology with the sequence shown in Seq ID No. 6.
53. DNA according to claim 52 wherein said homology is in the range of 65% to 99%.
15
54. DNA which hybridises to the DNA sequence shown in Seq ID No. 6 and which codes for at least part of *Spodoptera* ecdysone receptor.
55. DNA which is degenerate as a result of the genetic code to the DNA of any one of
20 claims 51 to 54.
56. DNA comprising part of the sequence shown in Seq ID No. 6, and which codes for at least part of the *Spodoptera* ecdysone receptor ligand binding domain.
25
57. DNA comprising a sequence which shows 60% or more homology with the sequence of claim 56.
58. DNA according to claim 57 wherein said homology is in the range of 65% to 99%.
30
59. DNA which hybridises to the DNA of any one of claims 56 to 58 and which codes for at least part of the *Spodoptera* ecdysone receptor ligand binding domain.
60. DNA which is degenerate as a result of the genetic code to the DNA of any one of
35 claims 56 to 58 and which codes for at least part of the *Spodoptera* ecdysone receptor ligand binding domain.

61. DNA comprising part of the sequence shown in Seq ID No. 6, and which codes for at least part of the *Spodoptera* ecdysone receptor hinge domain.
- 5 62. DNA comprising a sequence which shows 60% or more homology with the sequence of claim 61.
63. DNA according to claim 62 wherein said homology is in the range of 65% to 99%.
- 10 64. DNA which hybridises to the DNA of any one of claims 61 to 63 and which codes for at least part of the *Spodoptera* ecdysone receptor hinge domain.
- 15 65. DNA which is degenerate as a result of the genetic code to the DNA of any one of claims 61 to 63 and which codes for at least part of the *Spodoptera* ecdysone receptor hinge domain.
66. A polypeptide coded for by the DNA of any one of claims 51 to 65.
- 20 67. A fusion polypeptide comprising the polypeptide of claim 45 or 50 (when dependent upon claim 45) and functionally linked to a DNA binding domain and a transactivation domain.
68. Recombinant DNA comprising the DNA of any one of claim 8 to 14 functionally linked to DNA encoding a DNA binding domain and a transactivation domain.
- 25 69. A fusion polypeptide according to claim 67 or recombinant DNA according to claim 68 wherein the DNA binding domain and/or transactivation domain is fungal, bacterial, plant or mammalian.
- 30 70. A fusion polypeptide or recombinant DNA according to claim 69 wherein the DNA binding domain is GAL4 or A1CR/A.
71. A fusion polypeptide or recombinant DNA according to claim 69 or 70 wherein the transactivation domain is VP16.
- 35 72. A fusion polypeptide or recombinant DNA according to claim 69 wherein the DNA binding domain and/or transactivation domain is from a steroid receptor superfamily member.

73. A fusion polypeptide or recombinant DNA according to claim 72 wherein the DNA binding domain and/or transactivation domain is from a glucocorticoid or a *Spodoptera* ecdysone receptor.
- 5
74. A recombinant DNA construct comprising recombinant DNA of any one of claims 68 to 73; and DNA which codes for a gene operably linked to a promoter sequence and a hormone response element, which is responsive to the DNA binding domain coded for by said recombinant DNA.
- 10
75. A fusion polypeptide comprising the polypeptide of claim 46 or 50 (when dependent upon claim 46) and functionally linked to a ligand binding domain and a transactivation domain.
- 15 76. Recombinant DNA comprising the DNA of any of claims 15 to 21 functionally linked to DNA encoding a ligand binding domain and a transactivation domain.
77. A fusion polypeptide according to claim 75 or recombinant DNA according to claim 76 wherein the ligand binding domain and/or transactivation domain is fungal, bacterial, plant or mammalian.
- 20
78. A fusion polypeptide or recombinant DNA according to claim 77 wherein the transactivation domain is VP16.
- 25 79. A fusion polypeptide or recombinant DNA according to claim 77 wherein the ligand binding domain and/or transactivation domain is from a steroid receptor superfamily member.
80. A fusion polypeptide or recombinant DNA according to claim 79 wherein the ligand binding domain and/or transactivation domain is from a glucocorticoid or *Spodoptera* ecdysone receptor.
- 30
- 35 81. A recombinant DNA construct comprising recombinant DNA of any one of claims 76 to 80; and DNA which codes for a gene operably linked to a promoter sequence and a hormone response element, which is responsive to the DNA binding domain coded for by said recombinant DNA.

82. A fusion polypeptide comprising the polypeptide of claim 47 or 50 (when dependent upon claim 47) and functionally linked to a ligand binding domain and a DNA binding domain.
- 5 83. Recombinant DNA comprising the DNA of any one of claims 22 to 28 functionally linked to DNA encoding a ligand binding domain and a DNA binding domain.
- 10 84. A fusion polypeptide according to claim 82 or recombinant DNA according to claim 83 wherein the ligand binding domain and/or DNA binding domain is fungal, bacterial, plant or mammalian.
- 15 85. A fusion polypeptide or recombinant DNA according to claim 84 wherein the DNA binding domain is GAL4 or A1CR/A.
- 20 86. A fusion polypeptide or recombinant DNA according to claim 84 wherein the ligand binding domain and/or DNA binding domain is from a steroid receptor superfamily member.
- 25 87. A fusion polypeptide or recombinant DNA according to claim 86 wherein the ligand binding domain and/or DNA binding domain is from a glucocorticoid or *Spodoptera* ecdysone receptor.
- 30 88. A recombinant DNA construct comprising recombinant DNA of any one of claims 82 to 87; and DNA which codes for a gene operably linked to a promoter sequence and a hormone response element, which is responsive to the DNA binding domain coded for by said recombinant DNA.
- 35 89. A recombinant DNA construct comprising DNA according to any one of claims 1 to 7; and DNA comprising a sequence which codes for a gene operably linked to a promoter sequence and at least one hormone response element which is responsive to the DNA binding domain coded for by said DNA of any one of claim 1 to 7.
- 40 90. A recombinant DNA construct according to any one of claims 74, 81, 88 and 89 wherein said promoter sequence codes for a constitutive, spatially or temporally regulating promoter.

91. A recombinant DNA construct according to any one of claims 74, 81, 88 and 89 wherein there is more than one copy of the hormone response element.
92. A cell transformed with the DNA of any one of claims 1 to 42, and 51 to 65; the polypeptide of any one of claims 43 to 50; the fusion polypeptide of any one of claims 67, 70 to 73, 75, 77 to 80, 82 and 84 to 87; the recombinant nucleic acid of any one of claims 68 to 73, 76 to 80 and 85 to 87; or the recombinant DNA construct of any one of claims 74, 81, 88 and 89.
5
93. A cell according to claim 92 wherein said cell is a plant, fungal or mammalian cell.
10
94. A plant, fungus or mammal comprising the recombinant DNA construct of any one of claims 74, 81, 88 and 89.
95. A method of selecting compounds capable of being bound to an insect steroid receptor superfamily member comprising screening compounds for binding to said polypeptide of any one of claims 43 to 50 or the fusion polypeptide of any one of claims 67, 70 to 73, 75, 77 to 80, 82 and 84 to 87, and selecting said compounds exhibiting said binding.
15
96. A compound selected using the method of claim 95.
20
97. An agricultural or pharmaceutical composition comprising the compound of claim 96.
98. Use of the compound of claim 96 as an agrochemical or a pharmaceutical.
25
99. A method of producing a protein, peptide or polypeptide comprising introducing into the cell of claim 92, a compound which binds to the ligand binding domain in said cell.

1/56

Fig. 1.

Sequence ID 1

1 TGCG AGG GGT GCA AGG AGT TCT TCA GGC GGA GTG TAA CCA AAA ATG
ACGC TCC CCA CGT TCC TCA AGA AGT CCG CCT CAC ATT GGT TTT TAC

46 CAG TGT ACA TAT GCA AAT TCG GCC ATG CTT GCG AAA TGG ATA TGT
GTC ACA TGT ATA CGT TTA AGC CGG TAC GAA CGC TTT ACC TAT ACA

91 ATA TGC GGA GAA AAT GCC AAG AGT A
TAT ACG CCT CTT TTA CGG TGC TCA T

Fig. 2.

Fig.2 i.

361 CCT GCA GCA GAC ACA ACC CCT ACC TTC CAT GCC GTT ACC AAT GCC
 GGA CGT CTG TGT TGG GGA TGG AAG GTA CGG CAA TGG TTA CGG

406 ACC GAC AAC ACC CAA ATC AGA AAA CGA GTC AAT GTC ATC AGG TCG
 TGG CTG TTG TGG GTT TAG TCT TTT GCT CAG TTA CAG TAG TCC AGC

451 TGA GGA ACT GTC TCC AGC TTC GAG TGT AAA CGG CTG CAG CAC AGA
 ACT CCT TGA CAG AGG TCG AAG CTC ACA TTT GCC GAC GTC GTG TCT

496 TGG CGA GGC GAG GCG GCA GAA GAA AGG CCC AGC GCC GAG GCA GCA
 ACC GCT CCG CTC CGC CGT CTT CTT TCC GGG TCG CGG CTC CGT CGT

541 AGA AGA GCT ATG TCT TGT CTG CGG CGA CAG AGC CTC CGG ATA TCA
 TCT TCT CGA TAC AGA ACA GAC GAC GGC GCT GTC TCG GAG GCC TAT AGT

586 CTA CAA CGC GCT CAC ATG TGA AGG GTG TAA AGG ATT CTC CGG ATG CAA ATT TCC AAA GAA GTC CGC
 GAT GTT GCG CGA GTG TAC ACT TCC CAC ATT TCC AAA GAA GTC CGC

631 GAG TGT AAC CAA AAA TGC AGT GTA CAT ATG CAA ATT CGG CCA TGG CCG
 CTC ACA TTG GTT TTT ACG TCA CAT GTA TAC GTT TAA GCC GGT ACG

676 TTG CGA AAT GGA TAT CTA TAT GCG GAG AAA ATG TCA GGA GTG TCG
 AAC GCT TTA CCT ATA GAT ATA CGC CTC TTT TAC AGT CCT CAC AGC

721 GTT GAA GAA ATG TCT TGC GGT GGG CAT GAG GCC CGA GTG CGT GGT
 CAA CTT CTT TAC AGA ACG CCA CCC GTA CTC CGG GCT CAC GCA CGA

766 GCC GGA GAA CCA GTC GTG AAT GAA ACG GAA AGA AAA GGC GCA
 CGG CCT CTT GGT CAC ACG TTA CTT TGC CTT TCT CTT CTT CCG CGT

Fig.2 ii.

811 GAG GGA AAA AGA CAA ATT GCC CGT CAG TAC GAC GAC AGT AGA CGA
CTC CCT TTT TCT GTT TAA CGG GCA GTC ATG CTG CTG TCA TCT GCT

856 TCA CAT GCC TCC CAT GCA ATG TGA CCC TCC GCC CCC AGA GGC
AGT GTA CGG AGG GTA CGT TAC ACT GGG AGG CGG GGG TCT CCG

901 CGC TAG AAT TCT GGA ATG TGT GCA CGA GGT GGT GCC ACG ATT
GCG ATC TTA AGA CCT TAC ACA CGT GCT CCA CGA CGG TGC TAA

946 CCT GAA TGA GAA GCT AAT GGA ACA GAA CAG ATT GAA GAA CGT GCC
GGA CTT ACT CTT CGA TTA CCT TGT CGT GCT TAA CTT CGA CGA CGG

991 CCC CCT CAC TGC CAA TCA GAA GTC GTT GAT CGC AAG GCT CGT GTG
GGG GGA GTG ACG GTT AGT CTT CAG CAA CTA CGG TTC CGA GCA CAC

1036 GTA CCA CGA AGG CTA TGA ACA ACC TTC CGA GGA AGA CCT GAA GAG
CAT GGT CCT TCC GAT ACT TGT TGG AAG GCT CCT TCT GGA CGG CTT CTC

1081 GGT TAC ACA GTC GGA CGA CGA CGA AGA CTC GGA TAT GCC GTT
CCA ATG TGT CAG CCT GCT CCT GCT GCT TCT GAG CCT ATA CGG CAA

1126 CCG TCA GAT TAC CGA GAT GAC GAT TCT CAC AGT GCA GCT CAT CGT
GGC AGT CTA ATG GCT CTA CGT CTA AGA GTG TCA CGT CGA GTA GCA

1171 AGA ATT CGC TAA GGG CCT CCC GGG CTT CGC CAA GAT CTC GCA GTC
TCT TAA GCG ATT CCC GGA GGG CCC GAA GGG GTT CTA GAG CGT CAG

1216 GGA CCA GAT CAC GTT ATT AAA GGC GTG CTC AAG TGA CGT GAT GAT
CCT GGT CTA GTG CAA TAA TTT CCG CAC GAG TTC ACT CCA CTA CTA

1261 GCT CCG AGT GGC TCG GCG GTA TGA CGC GGC CAC CGA CAG CGT ACT
CGA GGC TCA CCG AGC CGC CAT ACT GCG CCG GTG GCT GTC CGA TGA

Fig.2 iii.

4/
56

1306	GTT CGC GAA CAA CCA GGC GTA CAC TCG CGA CAA CTA CCG CAA GGC CAA GCG CTT GGT CCG CAT GTG AGC GCT GAT GGC GTT CCG
1351	AGG CAT GGC GTA CGT CAT CGA GGA CCT GCT GCA CTT CTG TCG GTG TCC GTA CCG CAT GCA GTA GCT CCT GGA CGA CGT GAA GAC AGC CAC
1396	CAT GTA CTC CAT GAT GAT GGA TAA CGT GCA TTA TGC GCT GCT TAC GTA CAT GAG GTA CTA CTC CCT ATT GCA CGT AAT ACG CGA CGA ATG
1441	AGC CAT TGT CAT CTT CTC AGA CCG GCC CGG GCT TGA GCA ACC CCT TCG GTA ACA GTA GAA GAG TCT GGC CGG GCC CGA ACT CGT TGG GGA
1486	GTT GGT GGA GGA CAT CCA GAG ATA TTA CCT GAA CAC GCT ACG GGT CAA CCT CCT GTA GGT CTC TAT AAT GGA CCT GTG CGA TGC CCA
1531	GTA CAT CCT GAA CCA GAA CAG CGC GTC GCC CGG CGG CGT CAT CAT GTA GGA CCT GGT CTT GTC GCG CAG CGG GGC GCG GCA GTA
1576	CTT CGG CGA GAT CCT GGG CAT ACT GAC GGA GAT CGG CAC GCT GGG GAA GCC GCT CTA GGA CCC GTA TGA TGA CRG CCT CTA GGC GTG CGA CCC
1621	CAT GCA GAA CTC CAA CAT GTG CAT CTC CCT CAA GCT GAA GAA CAG GTA CGT CTT GAG GTT GTA CAC GTA GAG GGA GTT CGA CTT CTT GTC
1666	GAA GCT GCC GCC GTT CCT CGA GGA GAT CTG GGA CGT GGC GGA CGT CTT CGA CGG CGG CAA GGA GCT CCT CTA GAC CCT GCA CGG CCT GCA
1711	GGC GAC GAC GGC GAC GCC GGC GGT GGC GGA GGC GCC GCC GCC TCT CCG CTG CTG CCG CTG CCG CCA CCG CGG CCT CGG CGG CGG AGA

6/56

Fig.2 iv.

1756	AGC	CCC	CGC	CCC	GGC	CCC	GGC	CCC	GGC	CCC	GGC	CAC	CGT	CTA	GGG	CGC	
	TCG	GGG	GCA	GTG	GCA	GAT	GGC	GGG									
1801	CTC	AGG	AGA	GAA	CGC	TCA	TAG	ACT	GGC	TAG	TTT	TAG	TGA	AGT	GCA		
	GAG	TCC	TCT	CTT	GGC	AGT	ATC	TGA	CCG	ATC	AAA	ATC	ACT	TCA	CGT		
1846	CGG	ACA	CTG	ACG	TCG	ACG	TGA	TCA	ACC	TAT	TTA	TAA	GGG	CTG	CGA		
	GCC	TGT	GAC	TGC	AGC	TGC	ACT	AGT	TGG	ATA	AAT	ATT	CCT	GAC	GCT		
1891	ATT	TTA	CCA	CTT	AAG	AGG	GCA	CAC	CCG	TAC	CCG	ATT	TCG	TAC	GG		
	TAA	AAT	GGT	GAA	TTC	TCC	CGT	GTG	GGC	ATG	GGC	TAA	AGC	ATG	CC		

Total number of bases is: 1934.

7/56

Fig.3.
The sequence shown below is that of pSK16.1

Sequence ID3

3	9	15	21	27	33	39	45
1 CGC TGG TAT AAC AAC GGA CCA TTC CAG ACG CTC CGA ATG CTC GAG							
GGC ACC ATA TTG TTG CCT GGT AAG GTC TGC GAC GCT TAC GAG CTC							
<i>7/56</i>							
46	GAG AGC TCG TCT GAG GTG ACG TCG TCT TCA GCA CTG GGC CTG CCG	46	GAG AGC TCG TCT GAG GTG ACG TCG TCT TCA GCA CTG GGC CTG CCG	46	GAG AGC TCG TCT GAG GTG ACG TCG TCT TCA GCA CTG GGC CTG CCG	46	GAG AGC TCG TCT GAG GTG ACG TCG TCT TCA GCA CTG GGC CTG CCG
CTC TCG AGC AGA CTC CAC TGC AGC AGA AGT CGT GAC CCG GAC GGC							
<i>7/56</i>							
91	CCG GCT ATG GTG ATG TCC CCG GAA TCG CTC GCG TCG CCC GAG ATC	91	CCG GCT ATG GTG ATG TCC CCG GAA TCG CTC GCG TCG CCC GAG ATC	91	CCG GCT ATG GTG ATG TCC CCG GAA TCG CTC GCG TCG CCC GAG ATC	91	CCG GCT ATG GTG ATG TCC CCG GAA TCG CTC GCG TCG CCC GAG ATC
GGC CGA TAC CAC TAC AGG GGC CTT AGC GAG CGC AGC GGC CTC TAG							
<i>7/56</i>							
136	GGC CTG GAG CTG TGG GGC TAC GAC GAT GGC ATC ACT TAC AGC	136	GGC CTG GAG CTG TGG GGC TAC GAC GAT GGC ATC ACT TAC AGC	136	GGC CTG GAG CTG TGG GGC TAC GAC GAT GGC ATC ACT TAC AGC	136	GGC CTG GAG CTG TGG GGC TAC GAC GAT GGC ATC ACT TAC AGC
CCG CCG GAC CTC GAC ACC CCG ATG CTG CTA CCG TAG TGA ATG TCG							
<i>7/56</i>							
181	ATG GCA CAG TCG CTG GGC ACC TGC ACC ATG GAG CAG CAG CCC	181	ATG GCA CAG TCG CTG GGC ACC TGC ACC ATG GAG CAG CAG CCC	181	ATG GCA CAG TCG CTG GGC ACC TGC ACC ATG GAG CAG CAG CCC	181	ATG GCA CAG TCG CTG GGC ACC TGC ACC ATG GAG CAG CAG CCC
TAC CGT GTC AGC GAC CCG TGG ACG TGG TAC CTC GTC GTC GGG							

89
56

Fig.3 i.

226	CAG	CCG	CAG	CAG	CGG	CAG	CAG	ACA	CAA	CCC	CTA	CCT	TCC	ATG	
	GTC	GGC	GTC												
271	CCG	TTA	CCA	ATG	CCA	CCG	ACA	ACA	CCC	AAA	TCA	GAA	AAC	GAG	TCA
	GGC	AAT	GGT	TAC	GGT	GGC	TGT	TGT	GGG	TTT	AGT	CTT	TTG	CTC	AGT
316	ATG	TCA	TCA	GGT	CGT	GAG	GAA	CTG	TCT	CCA	GCT	TCG	AGT	GTA	AAC
	TAC	AGT	AGT	CCA	GCA	CTC	CTT	GAC	AGA	GGT	CGA	AGC	TCA	CAT	TTG
361	GGC	TGC	AGC	ACA	GAT	GGC	GAG	GGC	AGG	CGG	CAG	AAG	AAA	GGC	CCA
	CCG	ACG	TGC	TGT	CTA	CCG	CTC	CGC	TCC	GCC	GTC	TTC	TTT	CCG	GGT
406	GCG	CCG	AGG	CAG	CAA	GAA	GAG	CTA	TGT	CTT	GTC	TGC	GGC	GAC	AGA
	CGC	GGC	TCC	GTC	GTC	GTT	CTT	CTC	GAT	ACA	GAA	CAG	ACG	CCG	CTG
451	GCC	TCC	GGA	TAT	CAC	TAC	AAC	GCG	CTC	ACA	TGT	GAA	GGG	TGT	AAA
	CGG	AGG	CCT	ATA	GTG	ATG	TTG	GGC	GAG	TGT	ACA	CTT	CCC	ACA	TTT
496	GGT	TTC	TTC	AGG	CGG	AGT	GTA	ACC	AAA	ATT	GCA	GTG	TAC	ATA	TGC

9/56

CCA AAG AAG TCC GCC TCA CAT TGG TTT TTA CGT CAC ATG TAT ACG
 541 AAA TTC GGC CAT GCT TGC GAA ATG GAT ATC TAT ATG CGG AGA AAA
 TTT AAG CCG GTA CGA ACG CTT TAC CTA TAG ATA TAC GCC TCT TTT

 586 TGT CAG GAG TGT CGG TTG AAG AAA TGT CTT GCG GTG GGC ATG AGG
 ACA GTC CTC ACA GCC AAC TTC TTT ACA GAA CGC CAC CCG TAC TCC

 631 CCC GAG TGC GTG CCG GAG AAC CAG TGT GCA ATG AAA CGG AAA
 GGG CTC ACG CAC CAC GGC CTC TTG GTC ACA CGT TAC TTT GCC TTT

 676 GAG AAA AAG GCG CAG AGG GAA AAA GAC AAA TTG CCC GTC AGT ACG
 CTC TTT TTC CGC GTC TCC CTT TTT CTG TTT AAC GGG CAG TCA TGC

 721 ACG ACA GTA GAC GAT CAC ATG CCT CCC ATC ATG CAA TGT GAC CCT
 TGC TGT CAT CTG CTA GTG TAC GGA GGG TAG TAC CTT ACA CAC GTC GTG CTC

 766 CCC CCA GAG GCC GCT AGA ATT CTG GAA TGT GTG CAG CAC GAG
 GGC GGG GGT CTC CGG CGA TCT TAA GAC CTT ACA CAC GTC GTG CTC

 811 GTG GTG CCA CGA TTC CTG ATT GAG AAG CTA ATG GAA CAG AAC AGA
 CAC CAC GGT GCT AAG GAC TTA CTC TTC GAT TAC CTT GTC TGT TCT

 856 TTG AAG AAC GTG CCC CTC ACT GCC AAT CAG TCG TGT ATC
 AAC TTC TTG CAC GGG GGG GAG TGA CGG TTA GTC TTC AGC AAC TAG

 901 GCA AGG CTC GTG TGG TAC CAG GAA GGC TAT GAA CAA CCT TCC GAG
 CGT TCC GAG CAC ACC ATG GTC CTT CCG ATA CTT GTC AGG CTC

 946 GAA GAC CTG AAG AGG GTT ACA CAG TCG GAC GAC GAA GAC
 CTT CTG GAC TTC TCC CAA TGT GTC AGC CTC CTC AGC CTT CTC

Fig.3 ii.

Fig.3 iii.

991 TCG GAT ATG CCG TTC CGT CAG ATT ACC GAG ATG ACG ATT CTC ACA
 AGC CTA TAC GGC AAG GCA GTC TAA TGG CTC TAC TGC TAA GAG TGT
 1036 GTG CAG CTC ATC GTA GAA TTC GCT AAG GGC CTC CCG GGC TTC GCC
 CAC GTC GAG TAG CAT CTT AAG CGA TTC CCG GAG GGC CCG AAG CGG
 1081 AAG ATC TCG CAG TCG GAC CAG ATC ACG TTA AAG GCG TGC TCA
 TTC TAG AGC GTC AGC CTG GTC TAG TGC AAT AAT TTC CGC ACG AGT
 1126 AGT GAG GTG ATG ATG CTC CGA GTG GCT CGG CGG TAT GAC GCG GCG
 TCA CTC CAC TAC GAG GCT CAC CGA GCC GCC ATA CTG CGC CGG
 1171 ACC GAC AGC GTC CTG TTC GCG AAC AAC CAG GCG TAC ACT CGC GAC
 TGG CTG TCG CAT GAC AAG CGC TTG TTG GTC CGC ATG TGA GCG CTG
 1216 AAC TAC CGC AAG GCA GGC ATG GCG TAC GTC ATC GAG GAC CTG CTG
 TTG ATG GCG TTC CGT CCG TAC CGC ATG CAG TAG CTC CTG GAC GAC
 1261 CAC TTC TGT CGG TGC ATG TAC TCC ATG ATG ATG GAT AAC GTG CAT
 GTG AAG ACA GCC ACG TAC ATG AGG TAC TAC TAC CTA TTG CAC GTA
 1306 TAT GCG CTG CTT ACA GCC ATT GTC ATC TTC TCA GAC CGG CCC GGG
 ATA CGC GAC GAA TGT CGG TAA CAG TAG AAG AGT CTG GCC GGG CCC
 1351 CTT GAG CAA CCC CTG TTG GTG GAG GAC ATC CAG AGA TAT TAC CTG
 GAA CTC GTT GGG GAC AAC CAC CTC CTG TAG GTC TCT ATA ATG GAC
 1396 AAC ACG CTA CGG GTG TAC ATC CTG AAC CAG AAC AGC GCG TCG CCC
 TTG TGC GAT GCC CAC ATG TAG GAC TTG GTC TTG CGC AGC GGG
 1441 CGC GGC GCC GTC ATC TTC GGC GAG ATC CTG GGC ATA CTG ACG GAG

//56

Fig.3 iv.

	GGG CCG CAG TAG AAG CCG CTC TAG GAC CCG TAT GAC TGC CTC
1486	ATC CGC ACG CTG GGC ATG CAG AAC TCC AAC ATG TGC ATC TCC CTC TAG GCG TGC GAC CCG TAC GTG AGG TTG TAC ACG TAG AGG GAG
1531	AAG CTG AAC AGG AAG CTG CCG CCG TTC CTC GAG GAG ATC TGG TTC GAC TTC TTG TCC TTC GAC GGC GGC AAG GAG CTC CTC TAG ACC
1576	GAC GTG GGG GAC GTG GCG ACG ACG GCG ACG CCC GTG GCG GCG GAG CTG CAC CGC CTG CAC CGC TGC CGC CGC TGC GGC CAC CGC CGC CTC
1621	GCG CCG GCG CCTT CTA GCC CCC GCC CGG CCC GCC CGG CCG CCC GCC CGC GGC CGC GGA GAT CGG GGG CGG GGC GGG CGG GGC GGG CGG
1666	ACC GTC TAG CGC GCC CCTT CTA TCA GGA GAG AAC GCT CAT AGA CTG GCT AGT TGG CAG ATC CGG CGG AGT CCT CTC GTG ACT GCA GCT GCA CTA GTT GGA TAA
1711	TTT AGT GAA GTG CAC GGA CAC TGA CGT CGA CGT GAT CAA CCT ATT AAA TCA CTT CAC GTG CCT GTG ACT GCA GCT GCA CTA GTT GGA TAA
1756	TAT AAG GAC TGC GAA TTT TAC CAC TTA AGA GGG CAC ACC CGT ACC ATA TTC CTG ACG CTT AAA ATG GTG AAT TCT CCC GTG TGG GCA TGG
1801	CGA TTT CGT ACG TAT TCG GTG ACC GAC GAC GAT GCA GAG CGT GTG GCT AAA GCA TGC ATA AGC CAC TGG CTG CTC GCA CAC
1846	TAA TGT GAA TAT ATG TGT TGT TGA ACG ATT TGG AGA ATA TAT ATT ATT ACA CTT ATA TAC ACA ACA ACT TGC TAA ACC TCT TAT ATA TAA
1891	GCT GTT GCT GTT CGG GCC CGC ACC CCG TCG CCG GGC GGC GAT CCA CAA CGA CAA GCC CGG GCG TGC GGC AGC GGC CAG CCG CCG CTA

Fig.3 V.

12/56

1936	CGC	GGC	GCC	CGC	GGC	TTC	AGT	TTT	TCG	TTG	ACG	ACT	GAG	TTG	
	GGG	CCG	CCG	GGG	CCG	CGG	AAG	TCA	AAA	TAA	AGC	AAA	TGC	CTC	AAC
1981	GTC	ACT	CGG	ATA	CGA	CTG	TAT	GAT	AAG	ACT	TCG	TTC	GAT	AAG	TAC
	CAG	TGA	GGC	TAT	GCT	GAC	ATA	CTA	TTC	TGA	AGC	AAG	CTA	TTC	ATG
2026	ACC	TAC	TAA	ATT	ACA	CAT	ACG	TAC	GTA	GCT	TAC	GAG	AGT	TAT	TAG
	TGG	ATG	ATT	TAA	TGT	GTA	TGG	ATG	CAT	CGA	ATG	CTC	TCA	ATA	ATC
2071	AGA	CAA	AGA	ATA	TAA	GAA	GAA	GAT	GTT	TCT	ATT	GGG	TGA	AAA	GTT
	TCT	GTT	TCT	TAT	ATT	CTT	CTT	CTA	CAA	AGA	TAA	CCC	ACT	TTT	CAA
2116	GAT	AGT	TAT	GTT	TAT	TTA	CCA	AAA	TTA	ACA	ATA	CGT	TGA	TTA	
	CTA	TCA	ATA	CAA	ATA	AAT	GTT	TTT	AAT	TGT	TAT	TAT	GCA	ACT	AAT
2161	ACC	TTT	CGA	GTA	TAA	TAT	TGT	GAT	GAG	TCG	TCC	GCT	GTC	CAC	GTC
	TGG	AAA	GCT	CAT	ATT	ATA	ACA	CTA	CTC	AGC	AGG	CGA	CAG	GTG	CAG
2206	GCC	GTC	ACA	TGT	TTG	TTT	CTG	ATG	CAC	ACG	TGA	GGN	GCG	TTA	TCG
	CGG	CAG	TGT	ACA	AAC	AAA	GAC	TAC	GTG	TGC	ACT	CCN	CGC	AAT	AGC
2251	TGT	TTC	ATG	GTT	CCA	TCG	TCC	TGT	GCC	CGC	GAC	CCT	CGA	CTA	AAT
	ACA	AAG	TAC	CAA	GGT	AGC	AGG	ACA	CGG	GCG	CTG	GGA	GCT	GAT	TTA
2296	GAG	TAA	TTT	AAT	TTA	TTG	CTG	TGA	TTA	CAT	TTT	AAT	GTG	TTG	ATT
	CTC	ATT	AAA	TTA	AAT	AAC	GAC	ACT	AAT	GTA	AAA	TTA	CAC	AAC	TAA
2341	ATC	TAC	CAT	AGG	GTG	ATA	TAA	GTG	TGT	CTT	ATT	ACA	ATA	CAA	AGT
	TAG	ATG	GTA	TCC	CAC	TAT	ATT	CAC	ACA	GAA	TAA	TGT	TAT	GTT	TCA
2386	GTG	TGT	CGT	CGA	TAG	CTT	CCA	CAC	GAG	CAA	GCC	TTT	TGT	TTA	AGT

CAC	ACA	GCA	GCT	ATC	GAA	GGT	GTG	CTC	GTT	CGG	AAA	ACA	AAT	TCA
2431	GAT	TTA	CTG	ACA	TGG	ACA	CTC	GAC	CCG	GAA	CTT	C		
CTA	AAT	GAC	TGT	ACC	TGT	GAG	CTG	GGC	CTT	GAA	G			

Total number of bases is: 2464.

Fig.3 vi.

Fig. 4.

Sequence ID 4

Fig.4 i.

4/
56

190	200	210	220	230	240
AGTTAGTGGAGGAAAGTCAGTCAAAGCCTTCCTCGAGGATGTCCCTCGGGCTC					
M S L G A					
250	260	270	280	290	300
GTCGGATACCCGGAGCTTGACACGCCGCTCGCCGACATGAGACGCCGCTGGTATAACAACGGAC					
R G Y R C D T L A D M R R W Y N N G					
310	320	330	340	350	360
CATTCCAGACGGCTGCGAATGCTCGAGGAGGCTCGTGTAGGGTAGACGTCGTCTTCAGGCAC					
P F Q T L R M L E E S S S E V T S S S A					
370	380	390	400	410	420
TGGGCCTGCCGGCTATGGTGTATGTCGGGAATCGCTCGCGTCGCCGGAGATCGGGCG					
L G L P P A M V M S P E S L A S P E I G					

Fig.4 ii.

430 440 450 460 470 480
 |
 GCCTGGAGCTGTGGGCTACGACGATGGCATCACTTACAGCATGGCACAGTCGCTGGCA
 G L E L W G Y D D G I T Y S M A Q S L G

490 500 510 520 530 540
 |
 CCTGCACCATGGAGCAGGCCAGGCCAGGCCAGGCCAGCAGCAGCACACAACCCC
 T C T M E Q Q P Q P Q P Q Q P Q T Q P

550 560 570 580 590 600
 |
 TACCTTCCATGCCGTTACCAATGCCAACACCCAAATCAGAAAACGAGTCAAATGT
 L P S M P L P M P P T T P K S E N E S M

610 620 630 640 650 660
 |
 CATCAGGTCTCGTGAAGAACTGTCTCOAGCTTCGAGTGTAAACGGCTGCAGCACAGATGGCG
 S S G R E E L S P A S S V N G C S T D G

670 680 690 700 710 720
 |
 AGGGAGGGAGAAGAAGGCCAGGCCAGGGCAGGCCAGGAGGCTATGTCTTGTC
 E A R R Q K K G P A P R Q Q E E L C L V

/5
/6

Fig.4 iii.

730 740 750 760 770 780
 |
 GCGGCCGACAGAGCCTCCGGATATCACTACAACGGCTCACATGTGAAGGGTGTAAAGGTT
 C G D R A S G Y H Y N A L T C E G C K G

 790 800 810 820 830 840
 |
 TCTTCAGGGGGAGTGTAAACCAAAAATGCAGTGTACATATGCAAATTGGCCATGGCTTGCG

 F F R R S V T K N A V Y I C K F G H A C

 850 860 870 880 890 900
 |
 AAATGGATATCTATATGGGAGAAAATGTCAGGAGGTGTGGCTTGAGAAATGGCTTGCG

 E M D I Y M R R K C Q E C R L K K C L A

 910 920 930 940 950 960
 |
 TGGGCATGAGGCCGAGGTGCGTGGTGAAGAACCAAGTGTGCAAATGAAACGGAAAGAGA

 V G M R P E C V V P E N Q C A M K R K E

 970 980 990 1000 1010 1020
 |
 AAAAGGGCAGAGGGAAAAGACAAATTGCCGTAGTAGACGACAGTACGATCACA

 K K A Q R E K D K L P V S T T V D D H

Fig. 4 iv.

17/56

M	P	P	I	M	Q	C	D	P	P	P	P	E	A	R	I	L	E	C	
1030		1040			1050			1060			1070			1080					
TGCCTCCATCATGCAATGTGACCCCTCGCCCCAGGGCCCTAGAAATTCTGGAAATGTC																			
1090		1100			1110			1120			1130			1140					
TGCAGCACGAGGTGGTGCCACGATTCTGAATGAGAAGCTAAATGAAACAGAACAGATTGA																			
V	Q	H	E	V	V	P	R	F	L	N	E	K	L	M	E	Q	N	R	L
1150		1160			1170			1180			1190			1200					
AGAACGTGCCCTCACTGCAAATCAGAAGTCGTTGATCGCAAAGGCTCTCTGGTACCC																			
K	N	V	P	P	L	T	A	N	Q	K	S	L	I	A	R	L	V	W	Y
1210		1220			1230			1240			1250			1260					
AGGAAGGCTATGAAACAAACCTTCCGAGGAAGACCTGAAGAGGGTTACACAGTCGGACGAGG																			
Q	E	G	Y	E	Q	P	S	E	E	D	L	K	R	V	T	Q	S	D	E

Fig.4 v.

1270 1280 1290 1300 1310 1320
 | | | | | |
 ACGACGAAAGACTCGGATATGCCGGTTCAGATTACCGAGATGACGATTCTCACAGTGC

D D E D S D M P F R Q I T E M T I L T V

1330 1340 1350 1360 1370 1380
 | | | | | |
 AGCTCATCGTAGAATTCGCTAACGGGCCCTCCGGCTTCGCCAAGATCTGCAGTCGGACC

Q L I V E F A K G L P G F A K I S Q S D

1390 1400 1410 1420 1430 1440
 | | | | | |
 AGATCACGTTATTAAAGGCGTGCTCAAGTGAAGGTGATGATGCTCCGAGTGGCTGGCGGT

Q I T L L K A C S S E V M M L R V A R R

1450 1460 1470 1480 1490 1500
 | | | | | |
 ATGACCGGCCACCGACAGCGTACTGTTCGCGAACCAACCAGGGCTACACTCGCGACAACT

Y D A A T D S V L F A N N Q A Y T R D N

Fig.4 vi.

1510 1520 1530 1540 1550 1560
 | | | | |
 ACCGCAAGGCATGGCATCGGTACGGACCTGCTGCACTCTGTCGGTGCATGT
 Y R K A G M A Y V I E D L L H F C R C M

 1570 1580 1590 1600 1610 1620
 | | | | |
 ACTCCATGATGATGGATAACGTCATTATGCCATTGCTTACAGCCATTGTCATCTCTCAG
 Y S M M M D N V H Y A L L T A I V I F S

 1630 1640 1650 1660 1670 1680
 | | | | |
 ACCGGCGGGCTTGAGCAACCCCTGTGGAGGAGATCCAGAGATAATTACCTGAACA
 D R P G L E Q P L L V E E I Q R Y L N

 1690 1700 1710 1720 1730 1740
 | | | | |
 CGCTACGGGTGTACATCCTGAAACAGGGACGGCGTCCGGCGCCGCGTCATCTTCG
 T L R V Y I L N Q N S A S P R G A V I F

Fig.4 vii.

1750 1760 1770 1780 1790 1800
 |
 GCGAGATCCTGGCATACTGACGGAGATCCGCAACGCTGGCATGCCAAACATGT
 G E I L G I L T E I R T L G M Q N S N M

1810 1820 1830 1840 1850 1860
 |
 GCATCTCCCTCAAGCTGAAGAACAGGAAGCTGCCGGTTCCTCGAGGAGATCTGGGACG
 C I S L K L K N R K L P P F L E E I W D

1870 1880 1890 1900 1910 1920
 |
 TGGCGGACGTGGCGACGGGACGGCACGGCGACGCCGGTGGCGGGAGGGCGCCCTAGGCC
 V A D V A T T A T P V A A E A P A P L A

1930 1940 1950 1960 1970 1980
 |
 CGGCCCGCCGGCCGGCCACCGCTTAGCGGCCTCAGGAGAGAACGCTCAT
 P A P P A R P P A T V -

1990 2000 2010 2020 2030 2040
 |
 GACTGGCTAGTTAGTGAAGTGCACGGACACTGACGTCGACCGTGTACAAACCTATTATA

Fig.4 viii.

2050 |
 2060 |
 AGGACTGCGAATTACCACTTAAGAGGGCACACCCGTACCCGATTACGTATTTCGG
 2070 |
 2080 |
 TGACCGACCGATGCAGAGCGTGTGTAATGTGAATAATGTTGAAACGATTGG
 2090 |
 2100 |
 2110 |
 2120 |
 2130 |
 2140 |
 2150 |
 2160 |
 2170 |
 2180 |
 GAATATATTGGTGTGCTGGTTCGGCCGGCACGCCGTCGGCGGATCGCG
 2186
 2190 |
 2200 |
 2210 |
 2220 |
 2230 |
 2240 |
 GCGCCGGCTTCAGTATTTCGTTACGACTGAGTTGGTCACACTCGGATAACGACTGT
 2250 |
 2260 |
 2270 |
 2280 |
 2290 |
 2300 |
 ATGATAAGACTTCGTTCGATAAGTACACCTACTAAATTACACATAACGTACGTTACG
 2310 |
 2320 |
 2330 |
 2340 |
 2350 |
 2360 |
 AGAGTTATTAGAGACAAAGAATAAGAAGAGATGTTCTATTGGGTGAAAAGTTGATA

Fig.4 ix.

2410 2420 2430 2440 2450 2460
 | | | | | |
 GTTATGTTATTACCAAAATTAAACAATAAACGTTGATTAACCTTTCGAGTATAATTATT

 2470 2480 2490 2500 2510 2520
 | | | | | |
 GTGATGAGTCGGTCCGGCTGTCCACCGTCGCCGTCACATGTTCTGATGCACACGTCGAG

 2530 2540 2550 2560 2570 2580
 | | | | | |
 GNGCGTTATCGTGTTCATGGTTCCATCGTCCTGTGCCCGACCCCTCGACTAAATGAGT

 2590 2600 2610 2620 2630 2640
 | | | | | |
 AATTAAATTGCTGTGATTACATTAAATGTTGATTATCTACCATAAGGGTGATAT

 2650 2660 2670 2680 2690 2700
 | | | | | |
 AAGTGTCTTATTACAAATAAAAGTGTGTGTCGATAGCTCCACACGAGCAAGCCT

 2710 2720 2730 2740
 | | | |
 TTTGGTTAAGTGATTACTGACATGGACACTCGACCCCCGAACTC

Fig.5.

Sequence I.D. 5

BmECR	MRVENVDNVS	10
M _s ECR	-----	
H _v ECR	M-----	1
C _t ECR	-----	
A _a ECR	-----	
D _m ECR	-----	

23
/56

BmECR	FALNGRADEWCMHSVTRLDLIVREKSEVKAYVGGCPSVITDAGAYDALFD	60
M _s ECR	-----	
H _v ECR	-SLGARGYRRC-----	DTLAD
C _t ECR	-----	
A _a ECR	-----	
D _m ECR	-----	
BmECR	M-RRRWSNNNGGF-P-LRMLEESSSEVTSSSA-LGLPPAMVMSPESLASPEY	107
M _s ECR	M-RRRWSNNNGCFP-LRMFEESSEVTSSSA-FGMPAAMVMSPESLASPEY	47
H _v ECR	M-RRRWSNNNGFQTLMLEESSSEVTSSSA-LGLPPAMVMSPESLASPEI	64
C _t ECR	M-K-----TENLIVTT-VKVEPLNYASQSF	23
A _a ECR	MMKRRWSNNNGFTALRMLDDSSSEVTSSAAL----GMTMSPNSLGSPTY	46
D _m ECR	M-KRRWSNNNGF--MRLPEEESSEVTSSSNGLVLPSGVNMSPSSLDSHDY	47
 	 	*

Fig.5 i.
BmECR
MsECR
HvECR
CtECR
AaECR
DmECR

GALELW-----SY-----	114	114
GGLELW-----SY-----	55	55
GGLELW-----GY-----	72	72
GDMNI-----YGGAT-----	33	33
DELELW-SSYEDNAYNGHSV--LSNGNNN--LGGCGA-----	78	78
CDNDKWLCGNESGSFGSNGHGLSQQQSVITLAMHGCSSTLPAQTTIIP	97	97
BmECR-----DDGITY-----	121	121
MsECR-----DETMTN-----	61	61
HvECR-----DDGIT-----	77	77
CtECR-----KKQRILESDETMNH-----	46	46
AaECR-----ANNLIMNGTIVGNNNL-----NGMMN-----	98	98
DmECR-----INGNANGNGGSTNGQYVPGATNLGALLANGMLNGFTNGMQQQIQNGHGLIN	147	147
NTAQSLLGACACNMQQQLQP-----QQPHPAPPTLPTMP-----	154	154
YPAQSLLGACNAQQQQQQ-----QQQQPSAQPPLPSMP-----	94	94
YSMAQSLGCTMEEQQQQP-----QQQPQQTQPLPSMP-----	114	114
NQTNMNLESSNMNHNTIS-----GFSSPDVNYEAYSPNSKL-----DDGN-----	86	86
MASQAVQANANSIQHIVGN-----LINGVNPNQTLIPPLPS-----	134	134
STTPSTPTPLHQNLGGAGGGIGGGMGLHHANGTPNGLIGVVGGGG	197	197
BmECR-----LPMPPTPKSENESMSSGREELSPASSINGCSADA--D	190	190
MsECR-----LPMPPTPKSENESMSSGREELSPASSINGCSTDG--E	130	130
HvECR-----LPMPPTPKSENESMSSGREELSPASSVNGCSTDG--E	146	146
CtECR-----LDG-----K	98	98
AaECR-----IIQNTLMNTPRSESVNSISSSGREDLSPSSSLNGYT--DGSD	173	173
DmECR-----VGLGVGGGGVGGLGMQHTPRSDSVNSISSGRDDLSBPSSSLNGYSANESCD	247	247

Fig. 5

Fig.5 iii.

BmECR	VWYQEGYEQPSDEDLKRVTQTWQ-SDEEDEESDLPPFRQITEMTILTQQLI VMYQEGYEQPSEEDLKRVTQTWQLEEEEEEETDMPFRQITEMTILTQQLI VWYQEGYEQPSEEDLKRVTQS---DEDDEDSDMPPFRQITEMTILTQQLI IWYQDGYEQPSEEDLKRITTE--LEEEDQEHEANFRYITEVTILTQQLI IWYQDGYEQPSEEDLKRIMIG--SPNEEEEDQHDVHFRHITEITILTQQLI IWYQDGYEQPSEEDLRRIM-S--OPDENESOTDVSFHITEITILTQQLI	409 351 368 334 392 474
M _s ECR		
H _v ECR		
C _t ECR		
A _a ECR		
D _m ECR		

BmECR	VEFAKGLPGFSKISQSDQITLLKASSSEVMMRLVARRYDAASDSVLFANN VEFAKGLPGFSKISQSDQITLLKASSSEVMMRLVARRYDAATDSVLFANN VEFAKGLPGFAKISQSDQITLLKACSSSEVMMRLMARRYDHDSILFANN VEFAKGLPAFIKIPOEDQITLLKACSSSEVMMRLMARRYDHDSILFANN VEFAKGLPAFTKIPQEDQITLLKACSSSEVMMRLMARRYDHSSDIFANN VEFAKGLPAFTKIPQEDQITLLKACSSSEVMMRLMARRYDHSSDIFANN	459 401 418 384 442 524
M _s ECR	*****	
H _v ECR	*****	
C _t ECR	*****	
A _a ECR	*****	
D _m ECR	*****	

BmECR	KAYTRDNYRQGGMAYVIEDLILHFRCMFAMGMDNVHFALLTAIVIFSDRP QAYTRDNYRKAGMSYVIEDLILHFRCRCKMSMMDNVHYALLTAIVIFSDRP QAYTRDNYRKAGMAYVIEDLILHFRCRCKMSMMDNVHYALLTAIVIFSDRP TAYTKQTYQLAGMEETIDDLILHFRCRQMYALSIDNVEALLTAIVIFSDRP RSYTRDSYRMAGMADTIEDLILHFRCRQMFSLTVDNVEYALLTAIVIFSDRP RSYTRDSYRMAGMADNIEDLILHFRCRQMFSMKVDNVEYALLTAIVIFSDRP	509 451 468 434 492 574
M _s ECR	*****	
H _v ECR	*****	
C _t ECR	*****	
A _a ECR	*****	
D _m ECR	*****	

BmECR	GLEQPSLVEIQRYYLNTLRIYIINQNSASSRCAVIYGRILSVLTTELRTL GLEQPLLVVEIQRYYLNTLRIYIINQNSASPRCAVLFKGKILGVLTTELRTL GLEQPLLVVEDIQRYYLNTLRIYIINQNSASPRGAVIFGEILGILTEIRTL GLEKAEMVDIIQSYYTETLKVYIVRDHGGECSRCSVQFAKLIGILTEIRTL GLEQAELVEHIQSYYIDTLRIYIILNRHAGDPKCSVIFAKLSSILTELRTL GLEKAQLVEAIQSYYIDTLRIYIILNRHCGDMSMSLVFYAKLSSILTELRTL	559 501 518 484 542 624
M _s ECR	*****	
H _v ECR	*****	
C _t ECR	*****	
A _a ECR	*****	
D _m ECR	*****	

Fig.5 iv.

GTQNSNMCISLKLKNRKLPFFLEIWIWDAEVARR--	593
GTQNSNMCISLKLKNRKLPFFLEIWIWDAEVSTT--	535
GMQNSNMCISLKLKNRKLPFFLEIWIWDAEVSTT--	552
GNLNSEMCFSLKLRNRKLPRFLEEIVWDVGDNINQTTATTNTENIVRERIN	534
GNQNSEMCFSLKLKNRKLPRFLEEIVWDVQDIPPSMMAQMHSHTQSSS--	590
GNQNAEMCFSLKLKNRKLPKFLEEIVWDVHAIPPSVOSHLQITQEEDERLE	674
* * . * * . * * * . * * * . * * * . * * .	
BmECR	593
MSECR	535
HVECR	552
CtECR	536
AaECR	532
DmECR	724
BmECR	606
MSECR	556
HVECR	575
CtECR	536
AaECR	645
DmECR	774
BmECR	606
MSECR	556
HVECR	575
CtECR	536
AaECR	663
DmECR	824

28
56

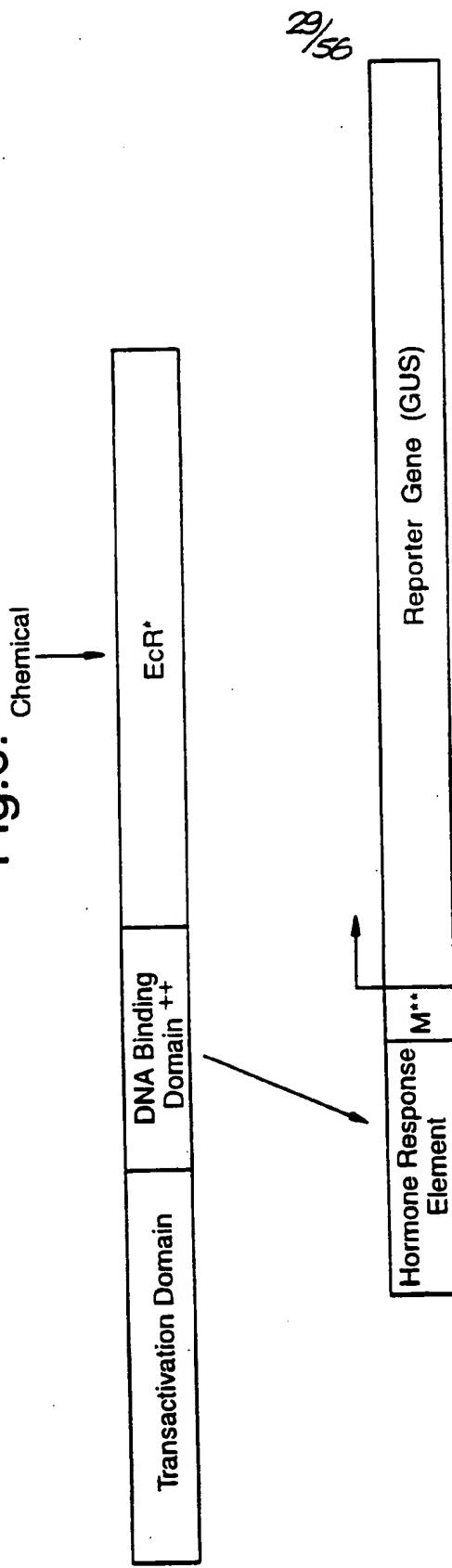
Fig.5 V.

BmECR	606	606
MSECR	556	556
HVECR	575	575
CtECR	536	536
AaECR	675	675
DmECR	874	874

-----GVVPGLGMLDQV-----PMGNGVGVGVGVGNGVSMYANAQTAMALMGVALHSHQQQLIGGVAVKSEH

BmECR	606	606
MSECR	556	556
HVECR	575	575
CtECR	536	536
AaECR	675	675
DmECR	878	878

Fig. 6.



++ Glucocorticoid receptor DNA binding and transactivation domains

* Insect ecdysone ligand binding domain

** Minimal 35S CaMV promoter

Fig.7.

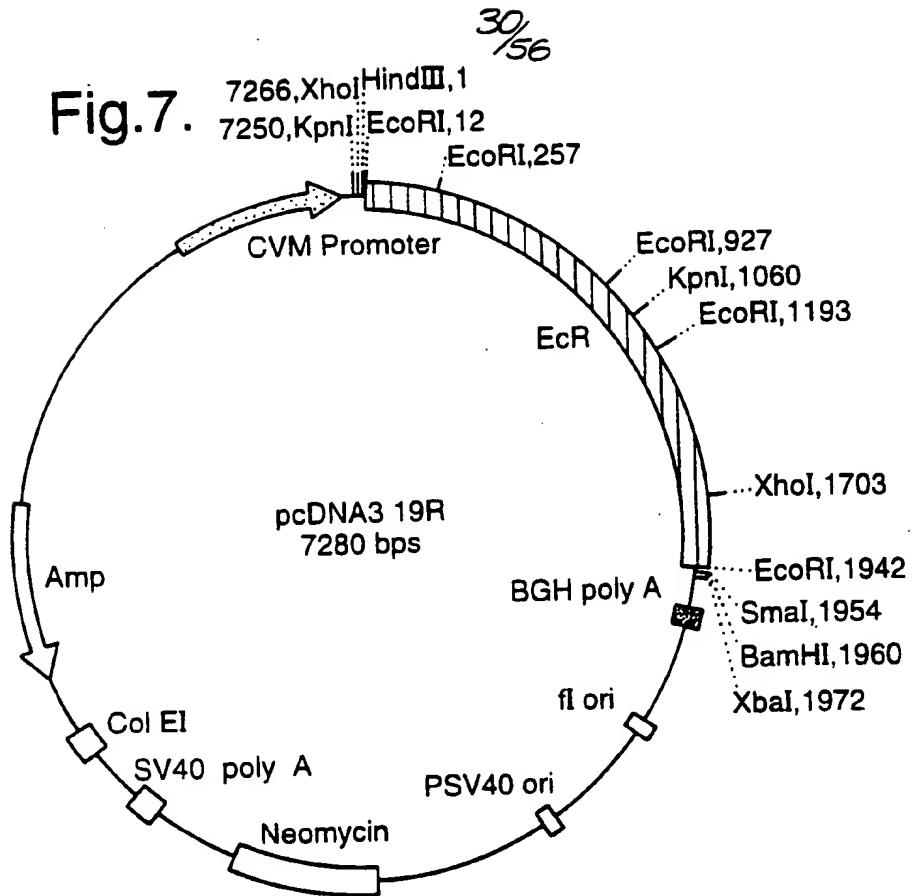
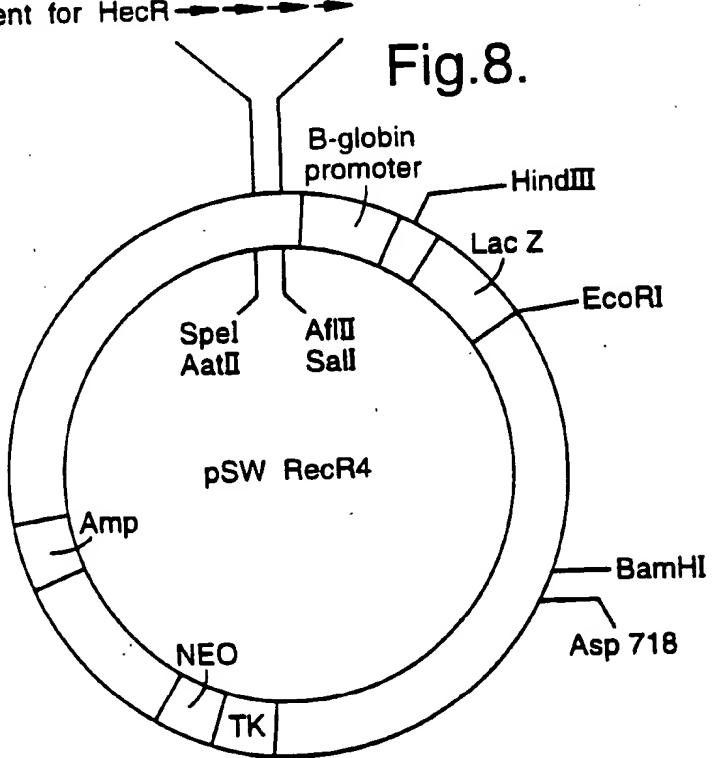


Fig.8.



31/56

Fig. 9.

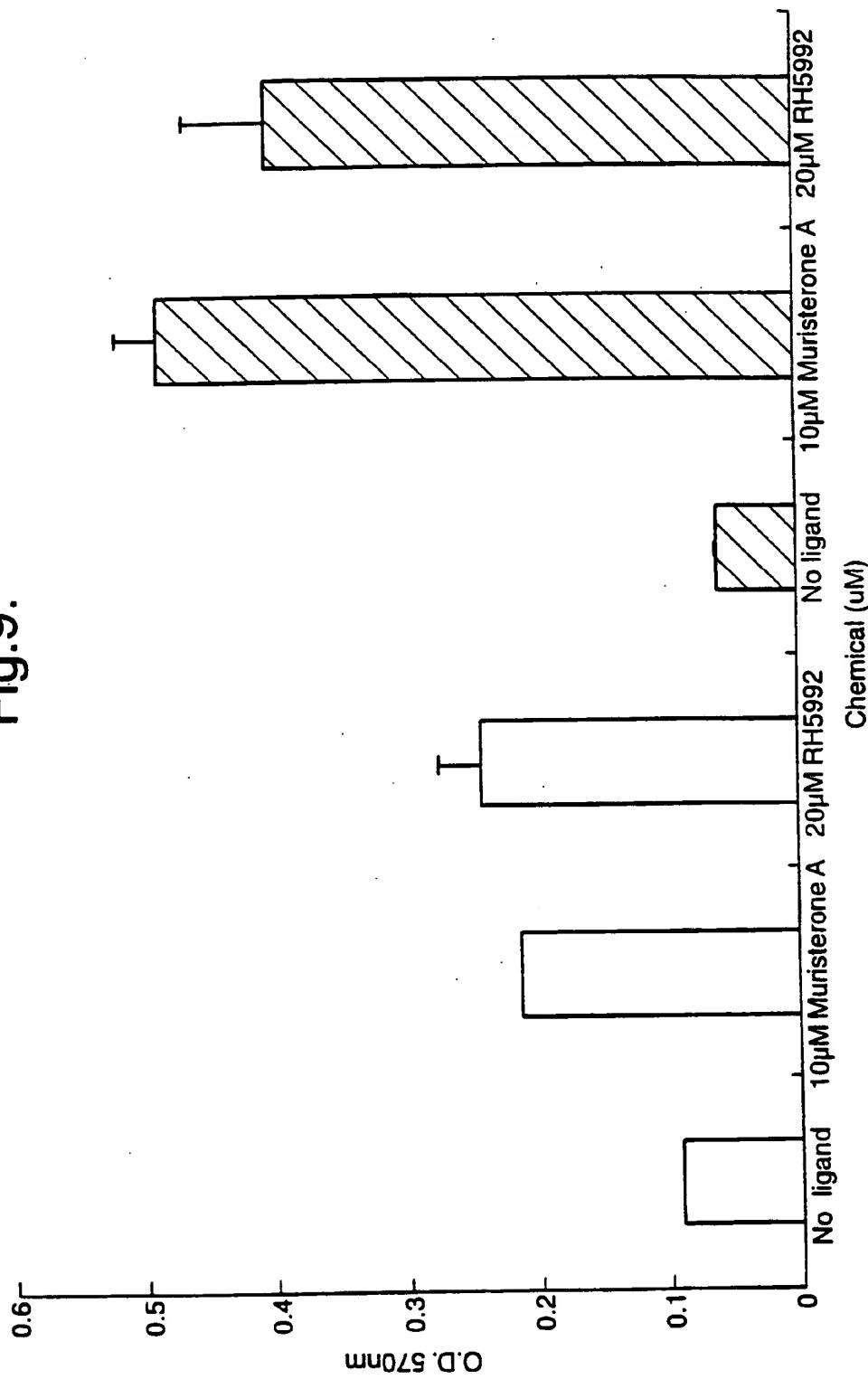
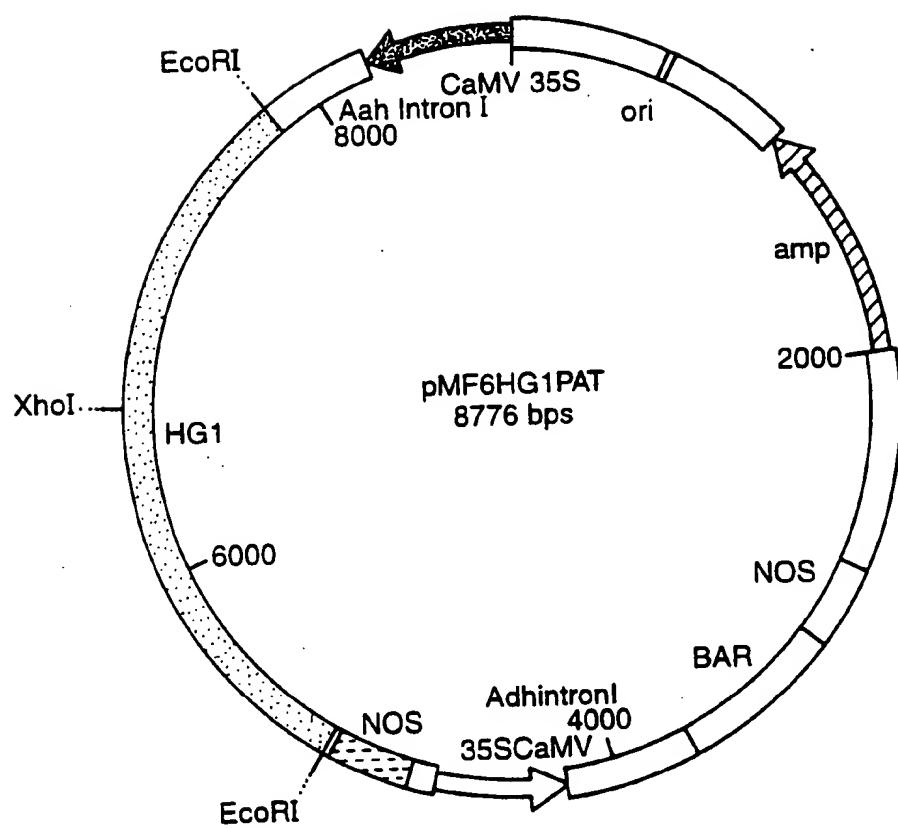
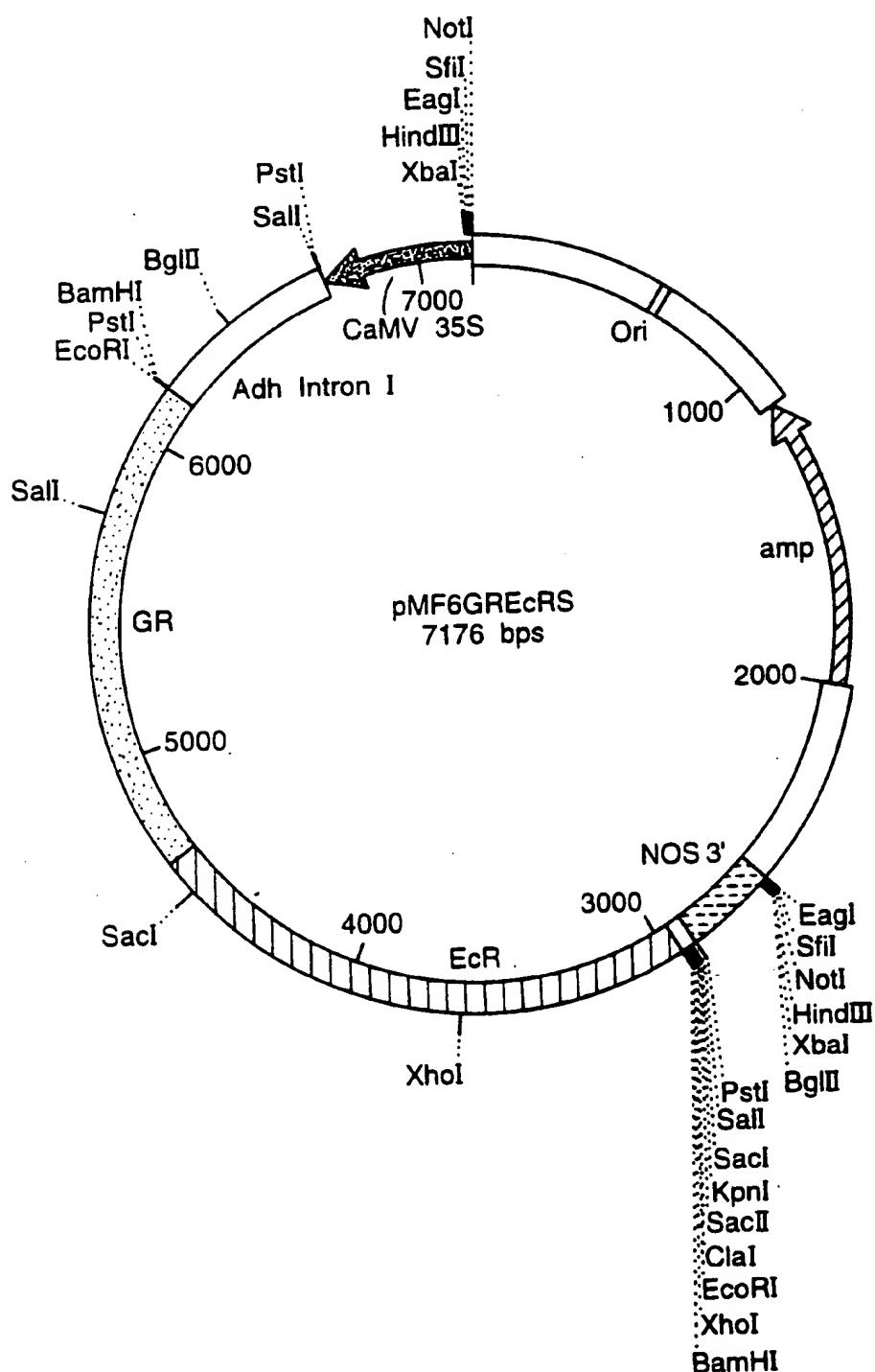


Fig.10.



33
56

Fig.11.



34/
55

Fig.12.

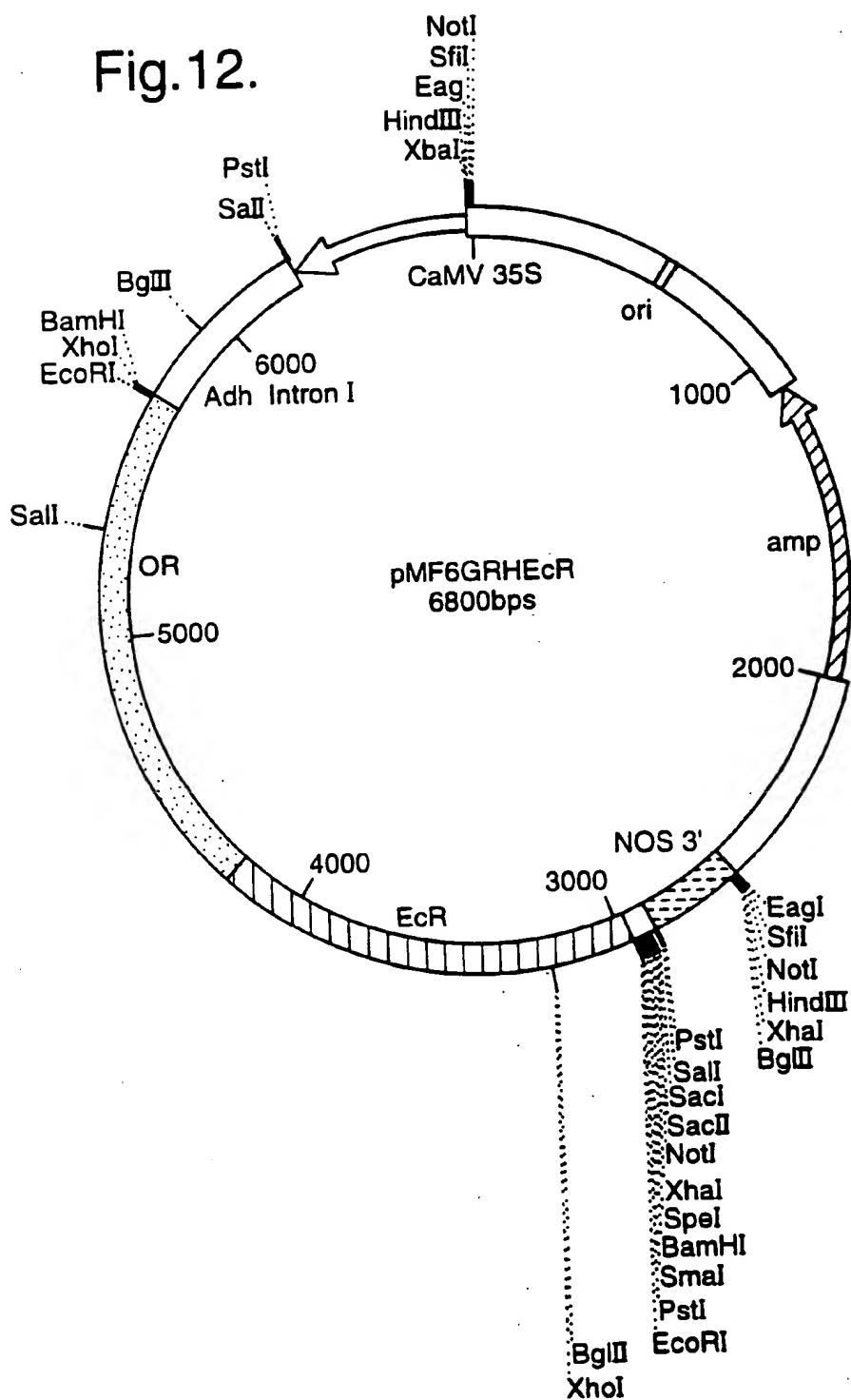


Fig.13.

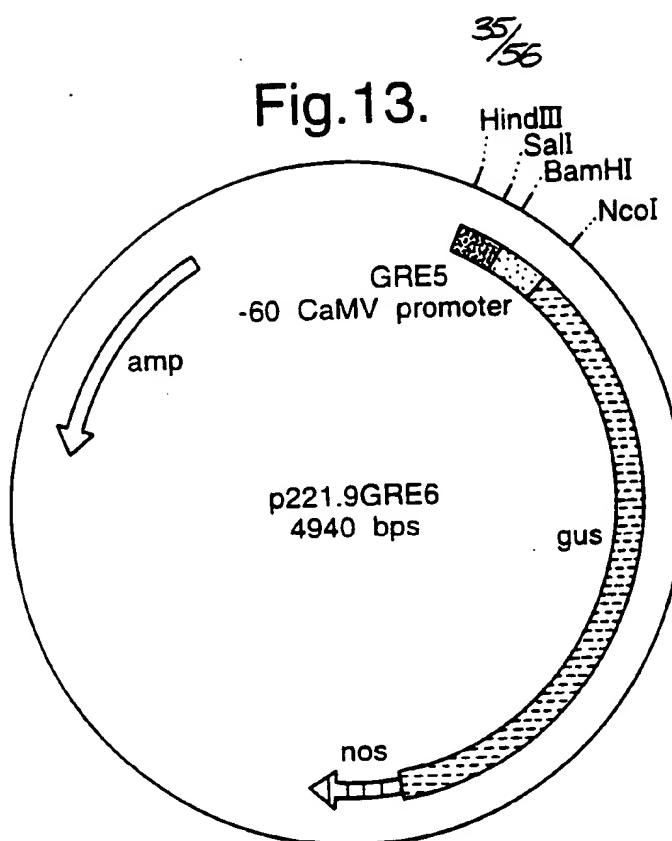
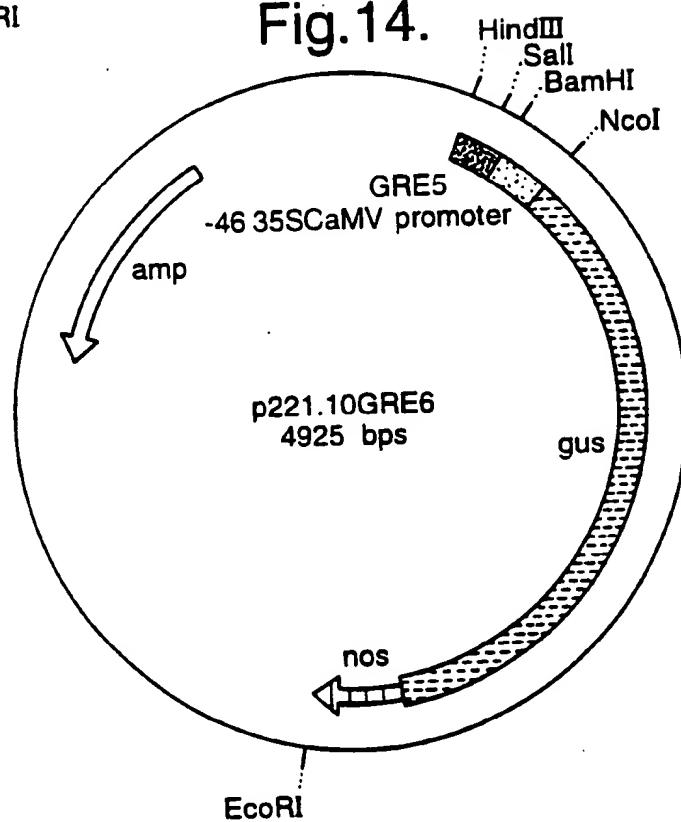


Fig.14.



36/56

Fig.15.

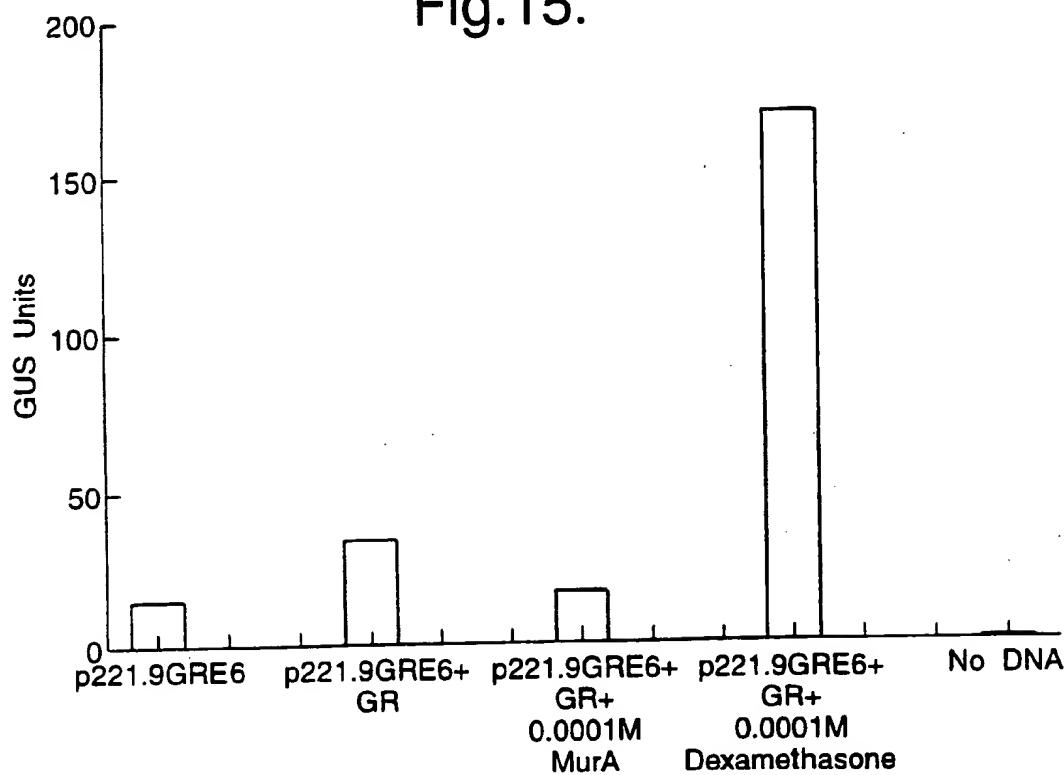
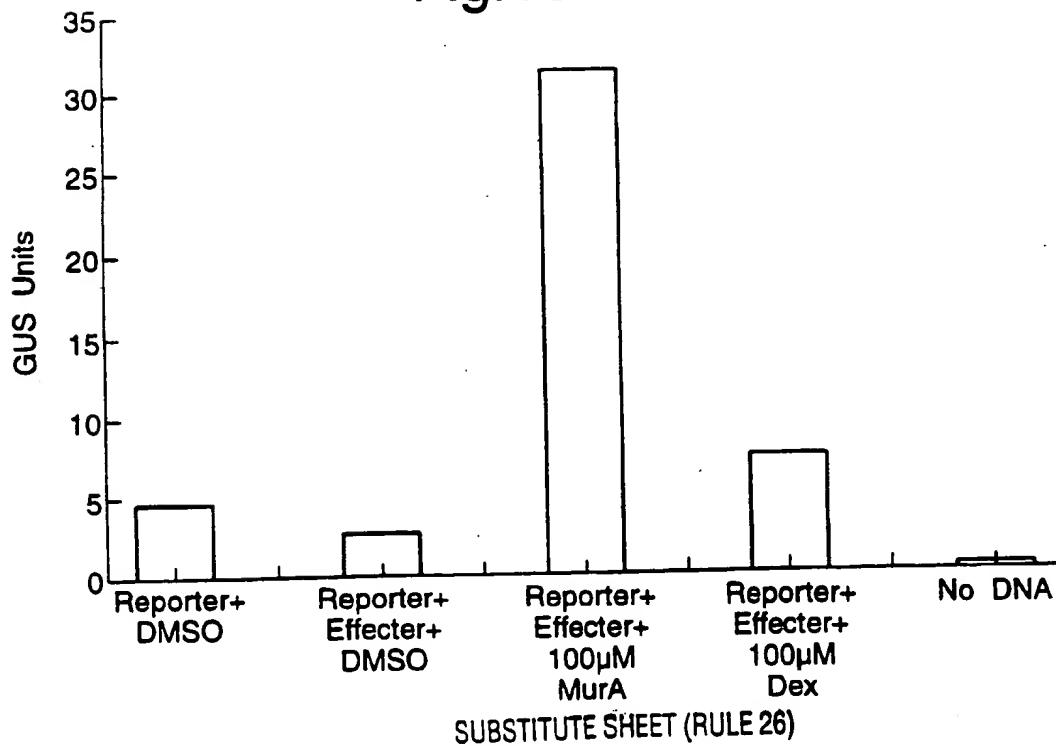


Fig.16.



SUBSTITUTE SHEET (RULE 26)

37/56

Fig.17.

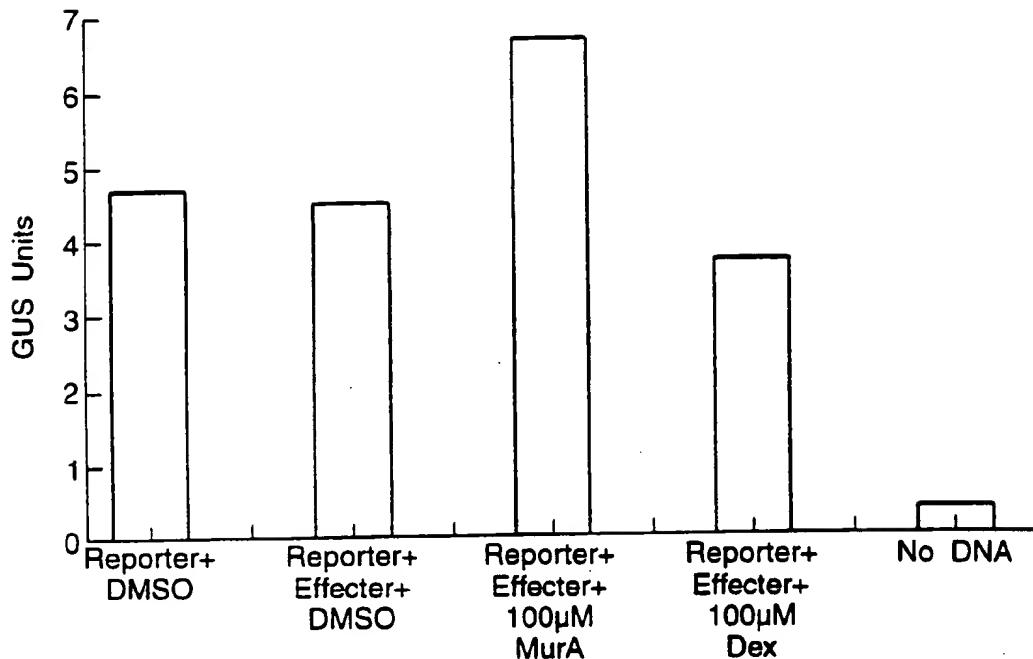
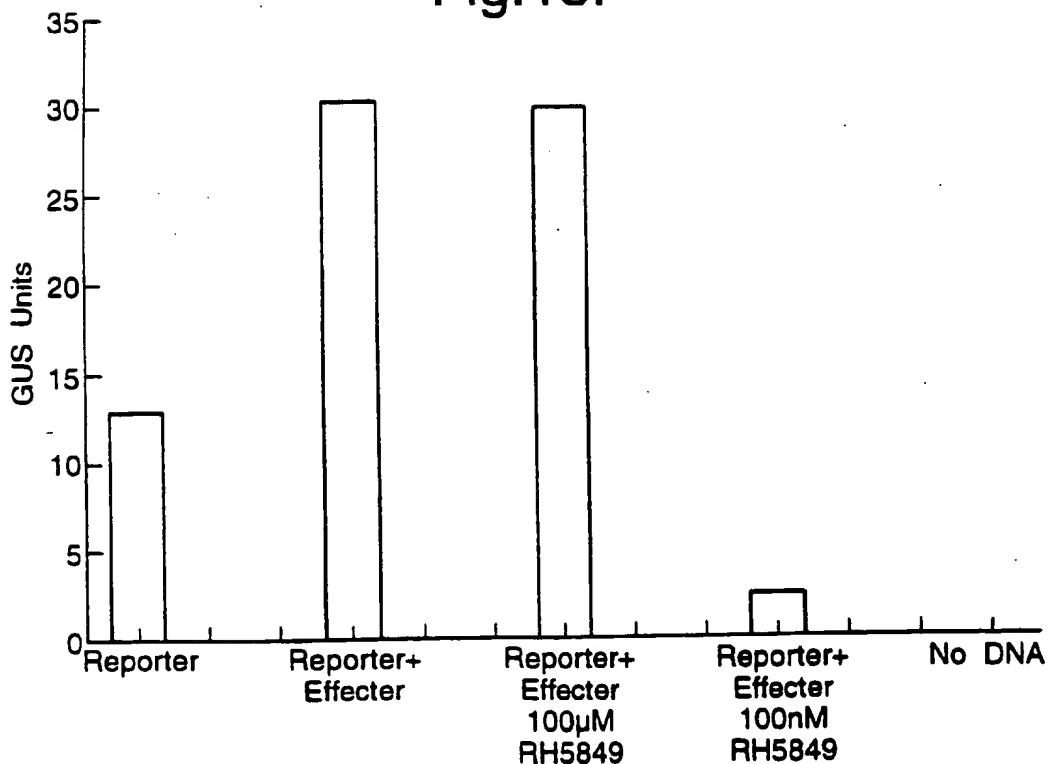


Fig.18.



38/
56

Fig.19.

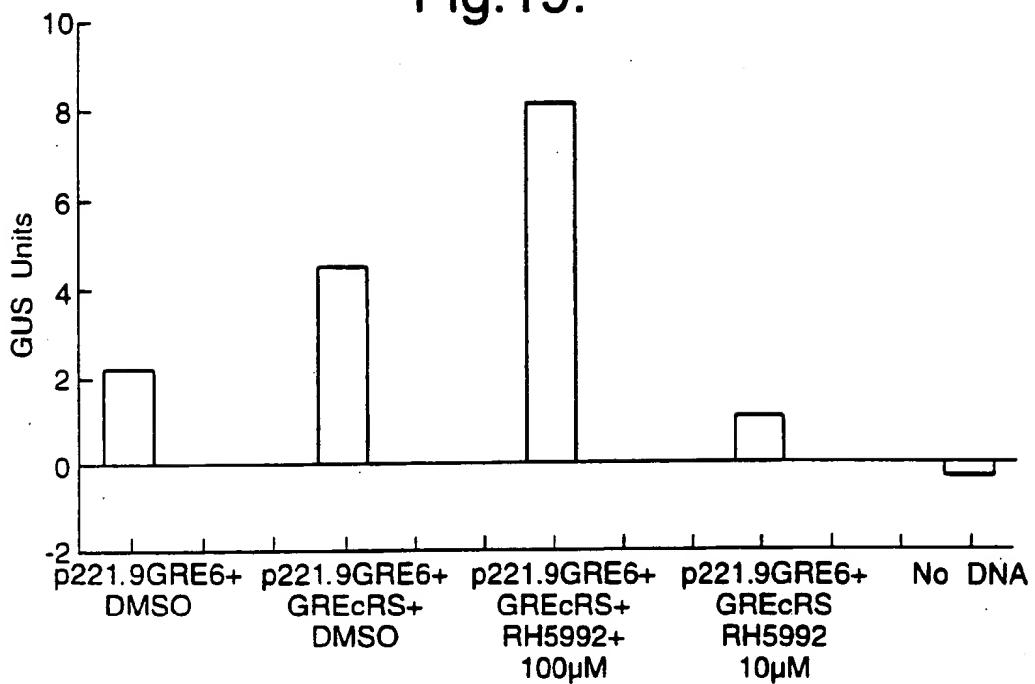
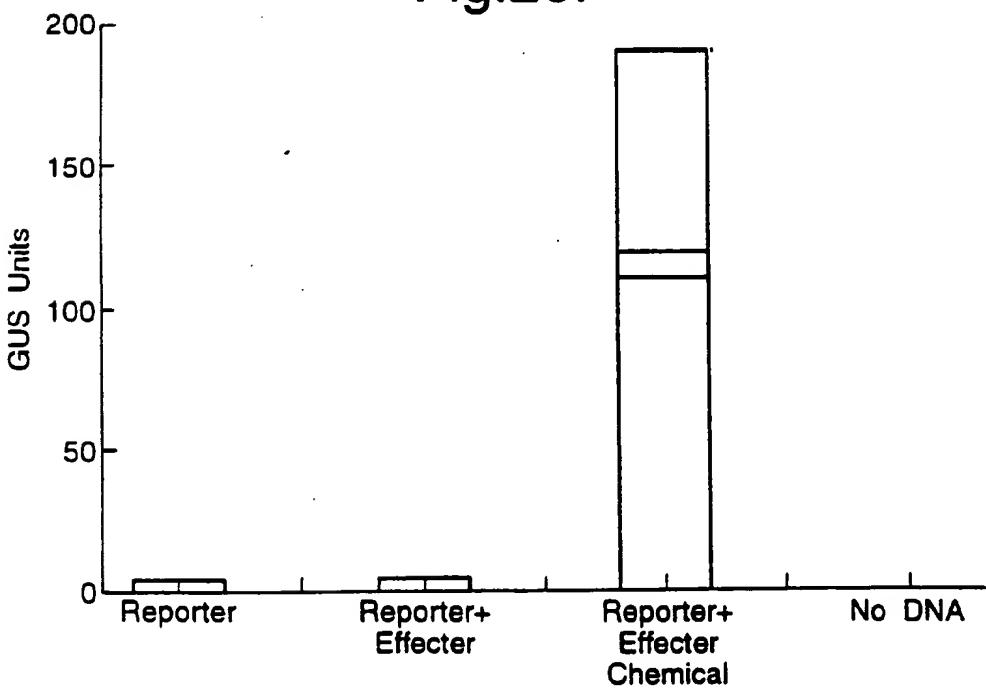
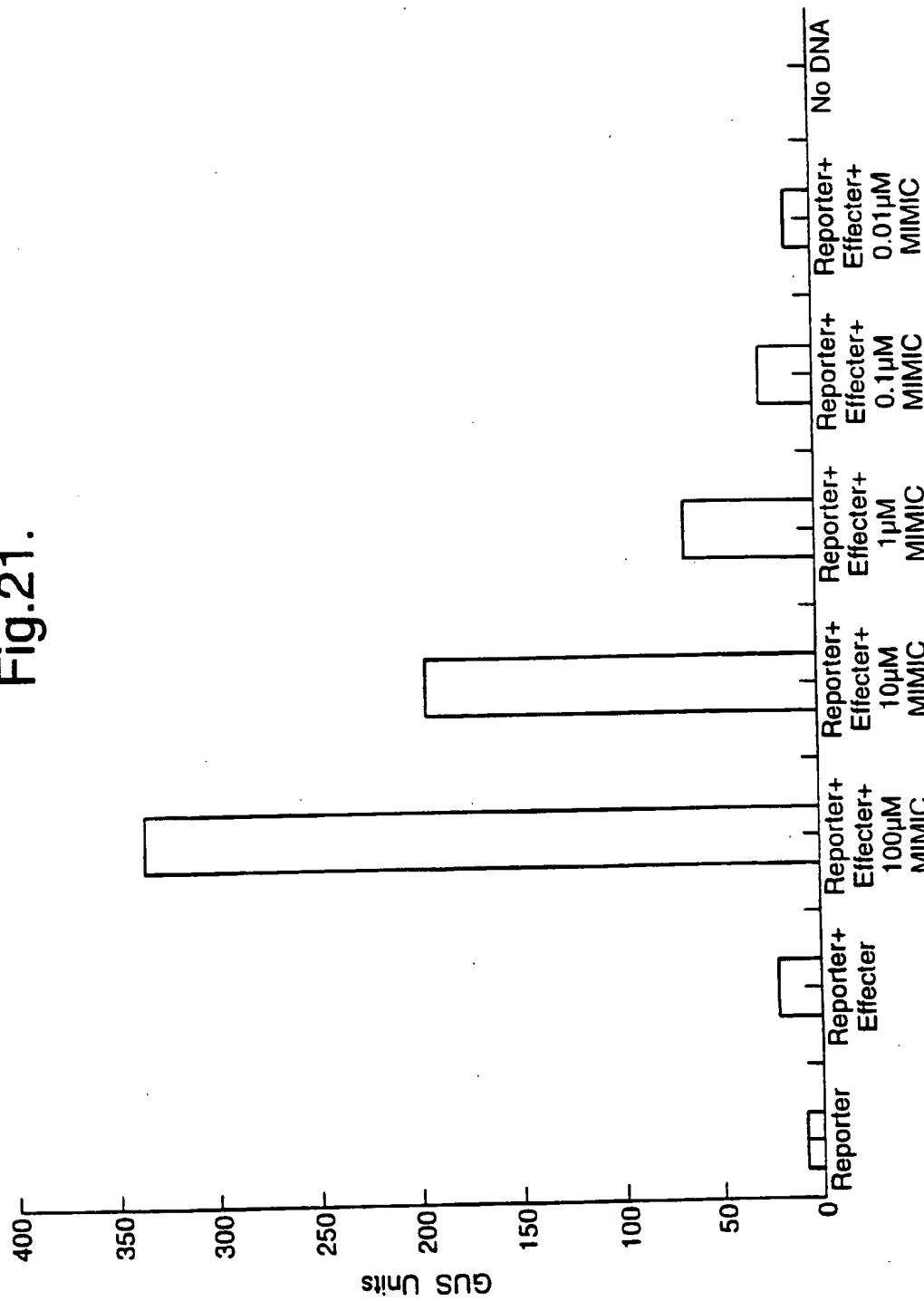


Fig.20.



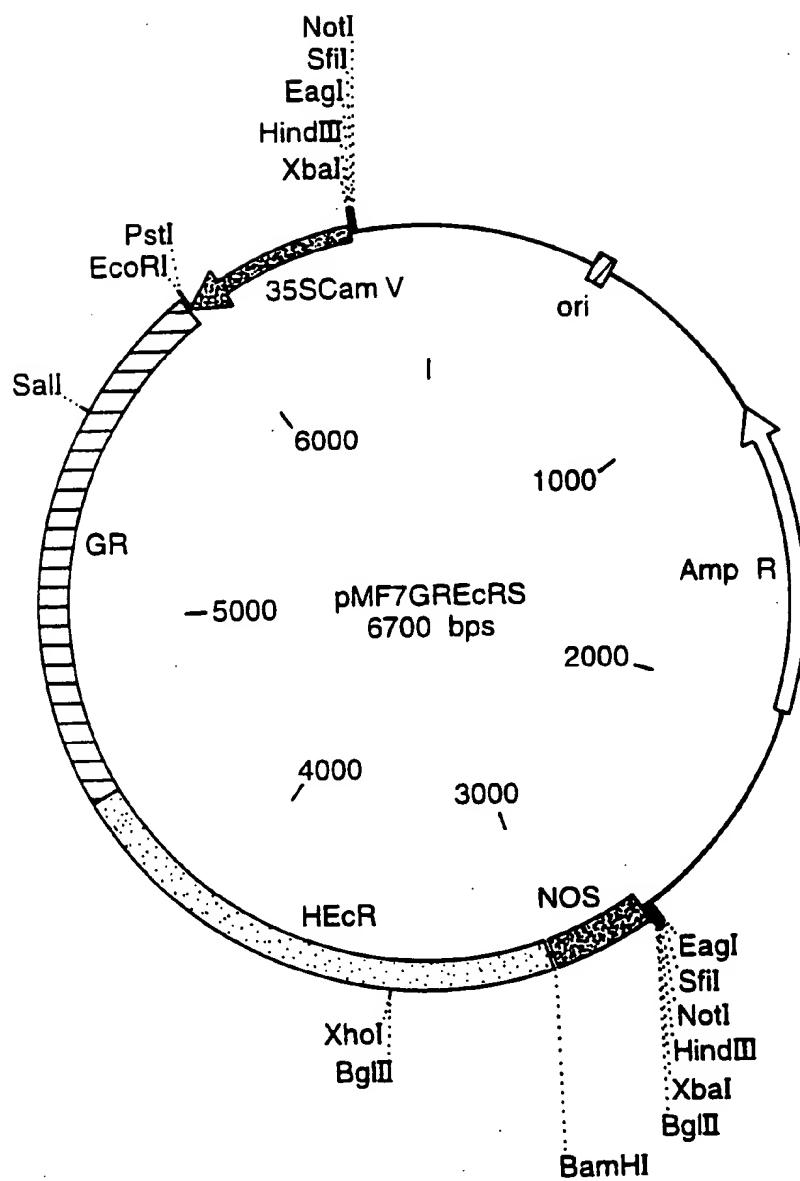
39
56

Fig. 21.



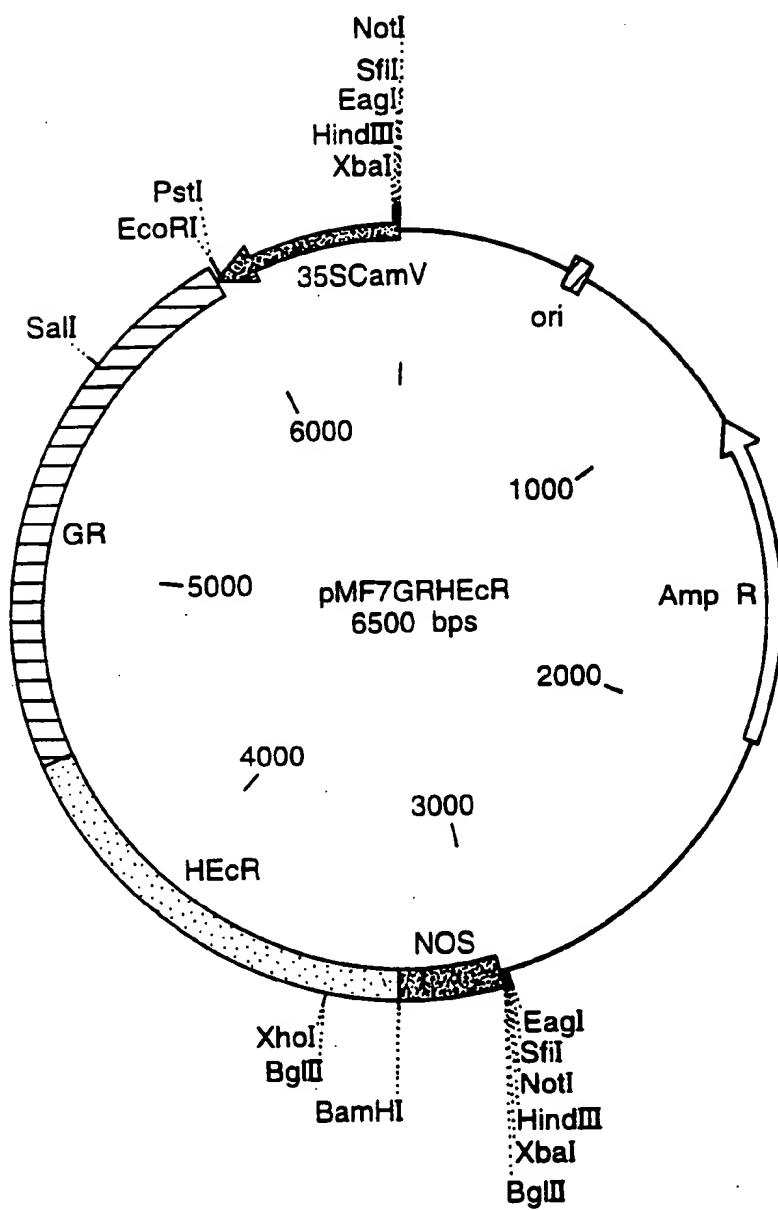
40/56

Fig.22.

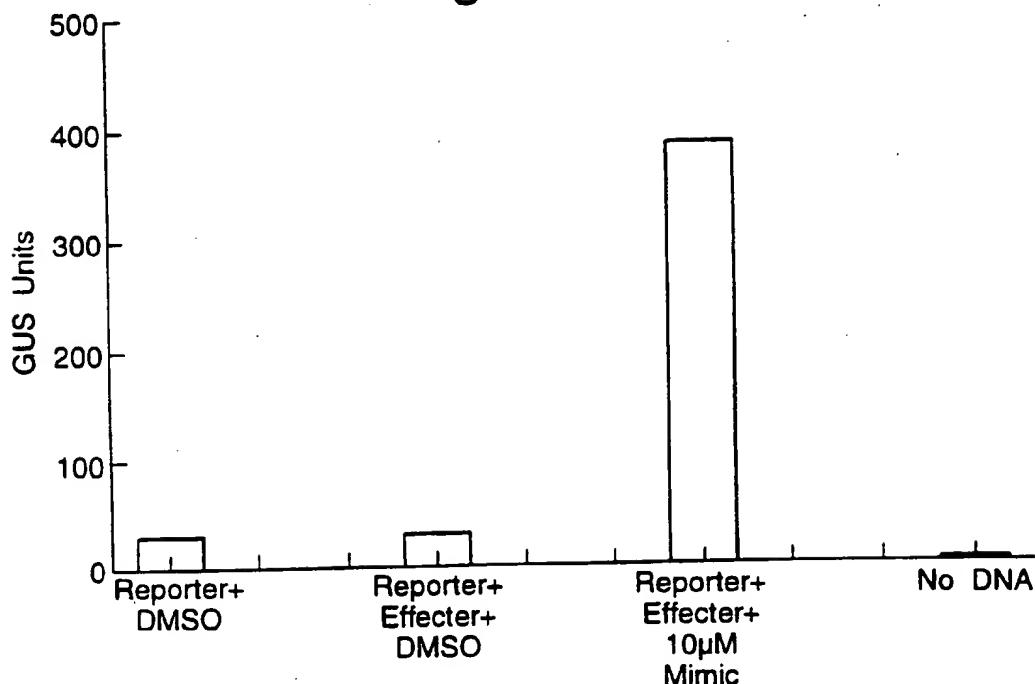
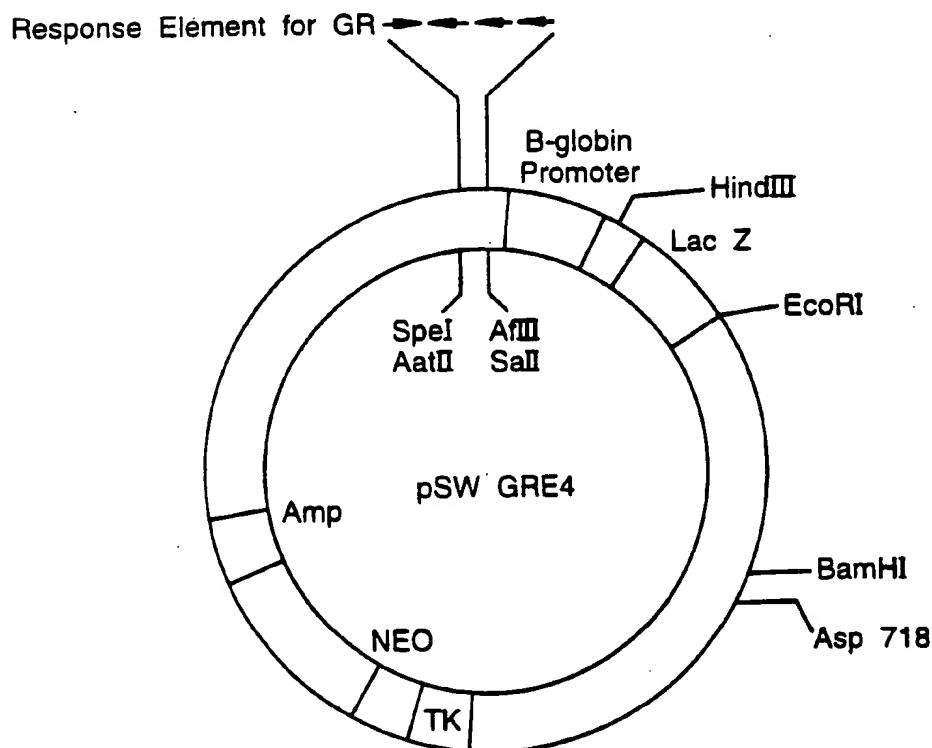


41/56

Fig.23.

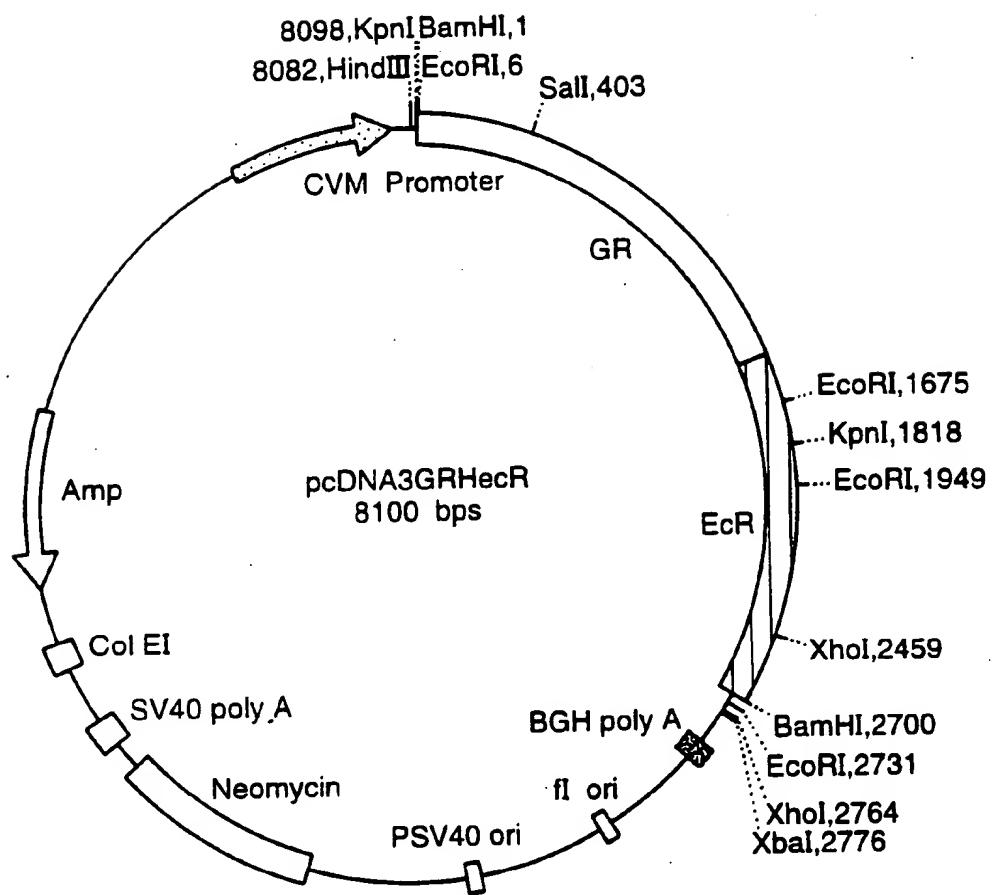


42
56
Fig.24.

**Fig.26.**

43/
56

Fig.25.



44/
56

Fig.27.

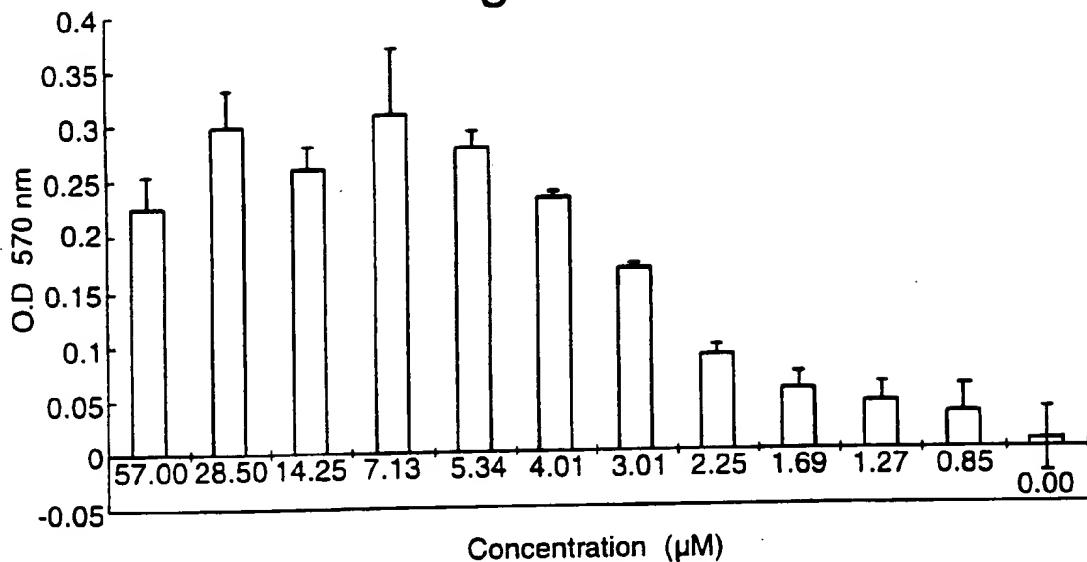
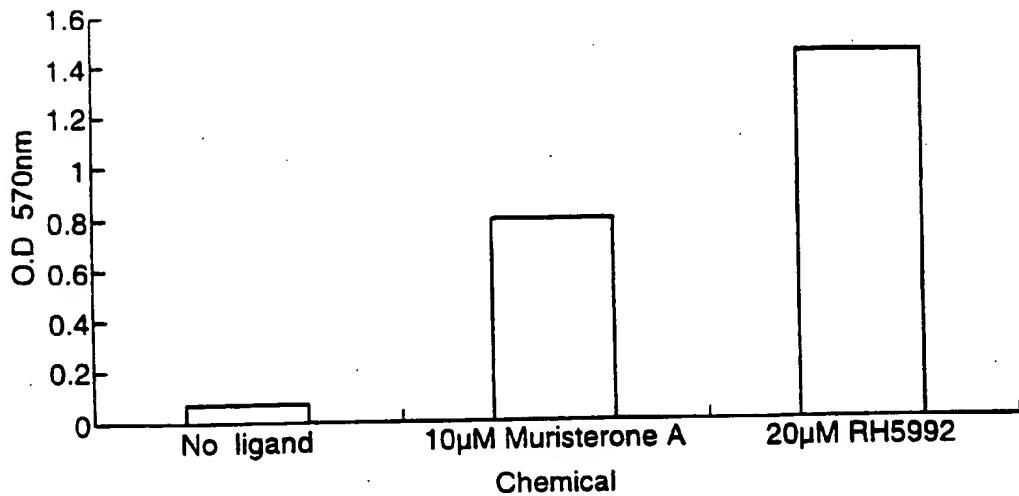


Fig.28.



45
56

Fig.29.

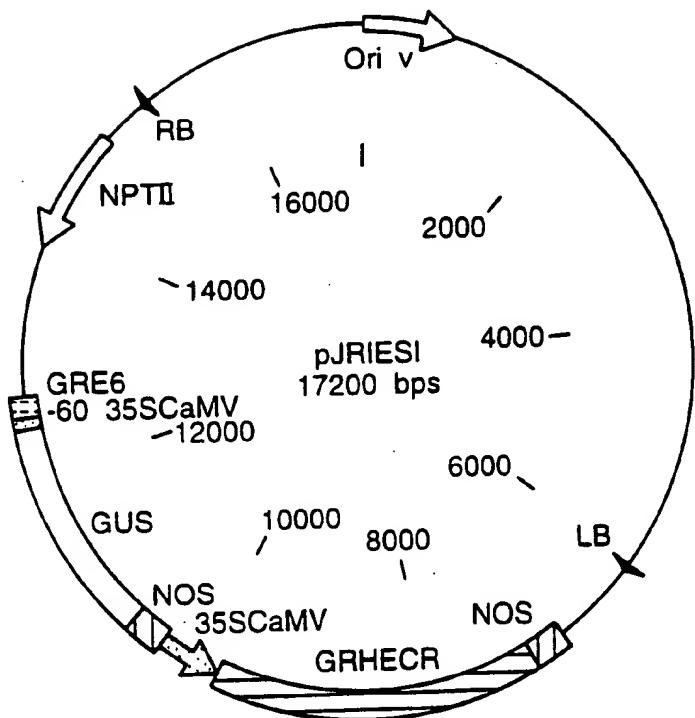
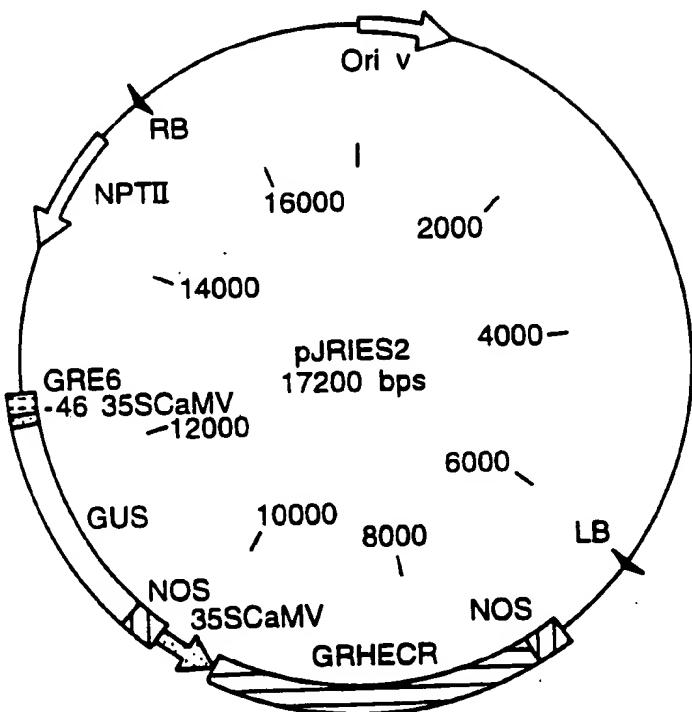


Fig.30.



46/56

Fig.31.

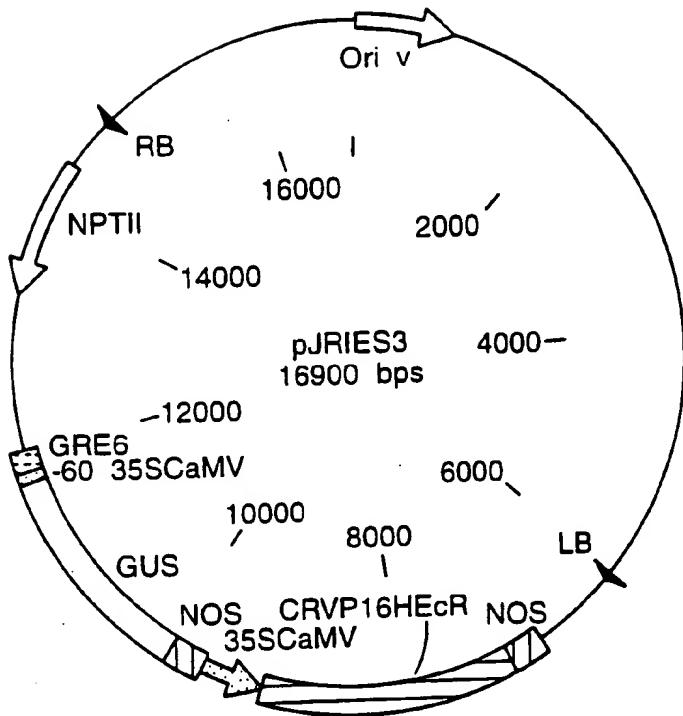


Fig.32.

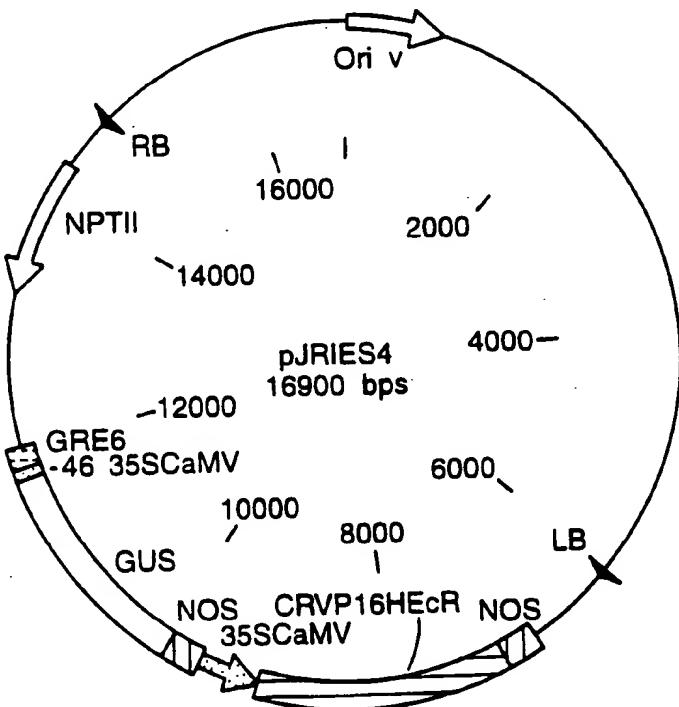


Fig.33.

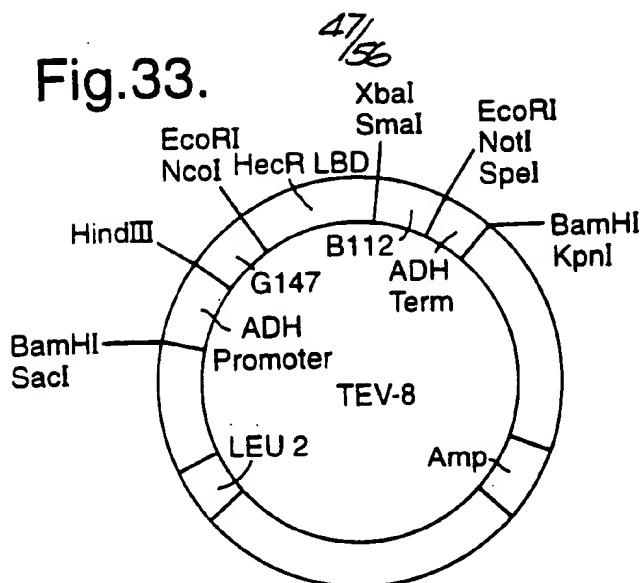


Fig.34.

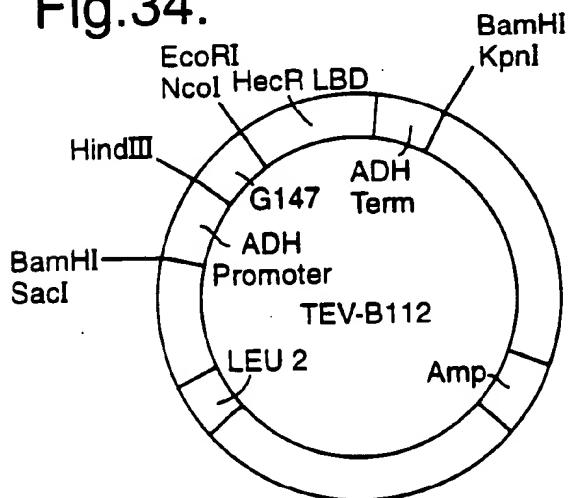
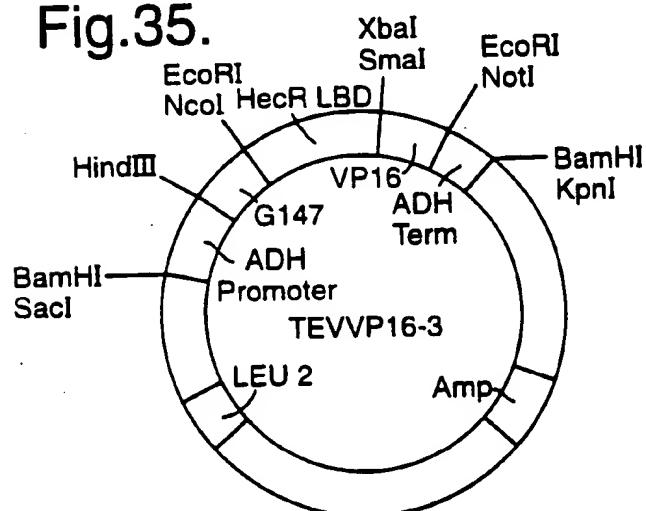
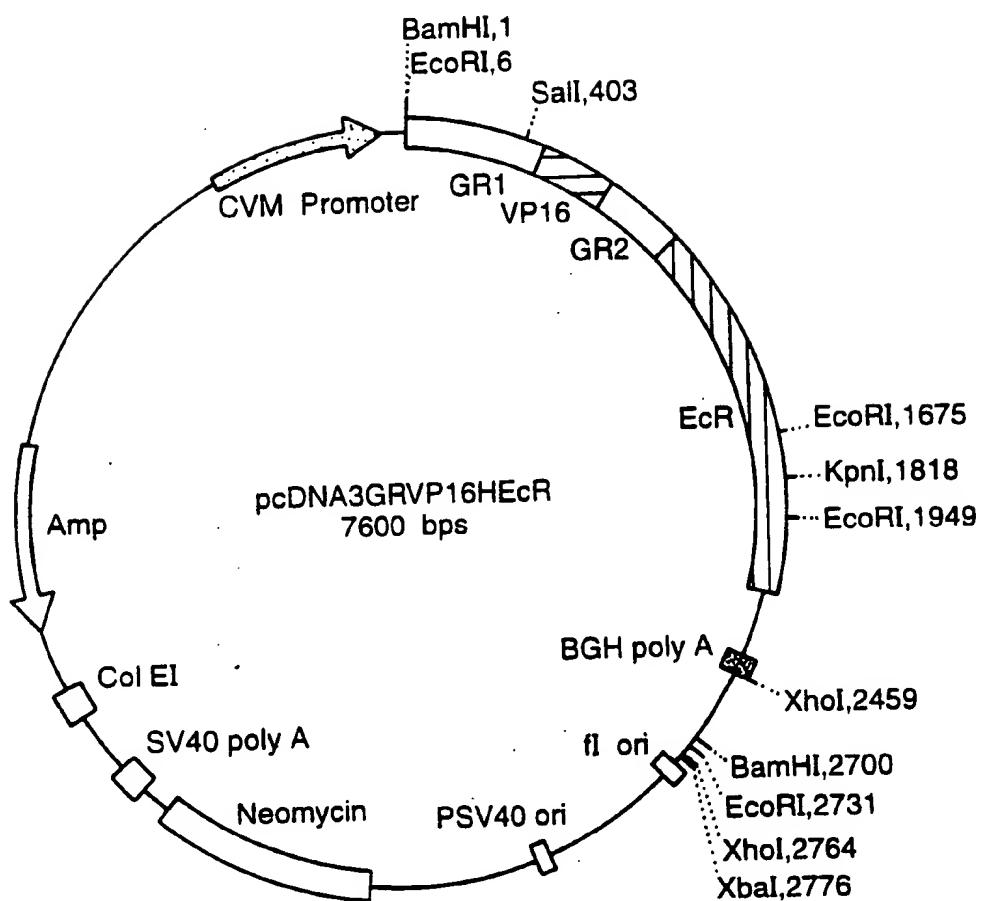


Fig.35.



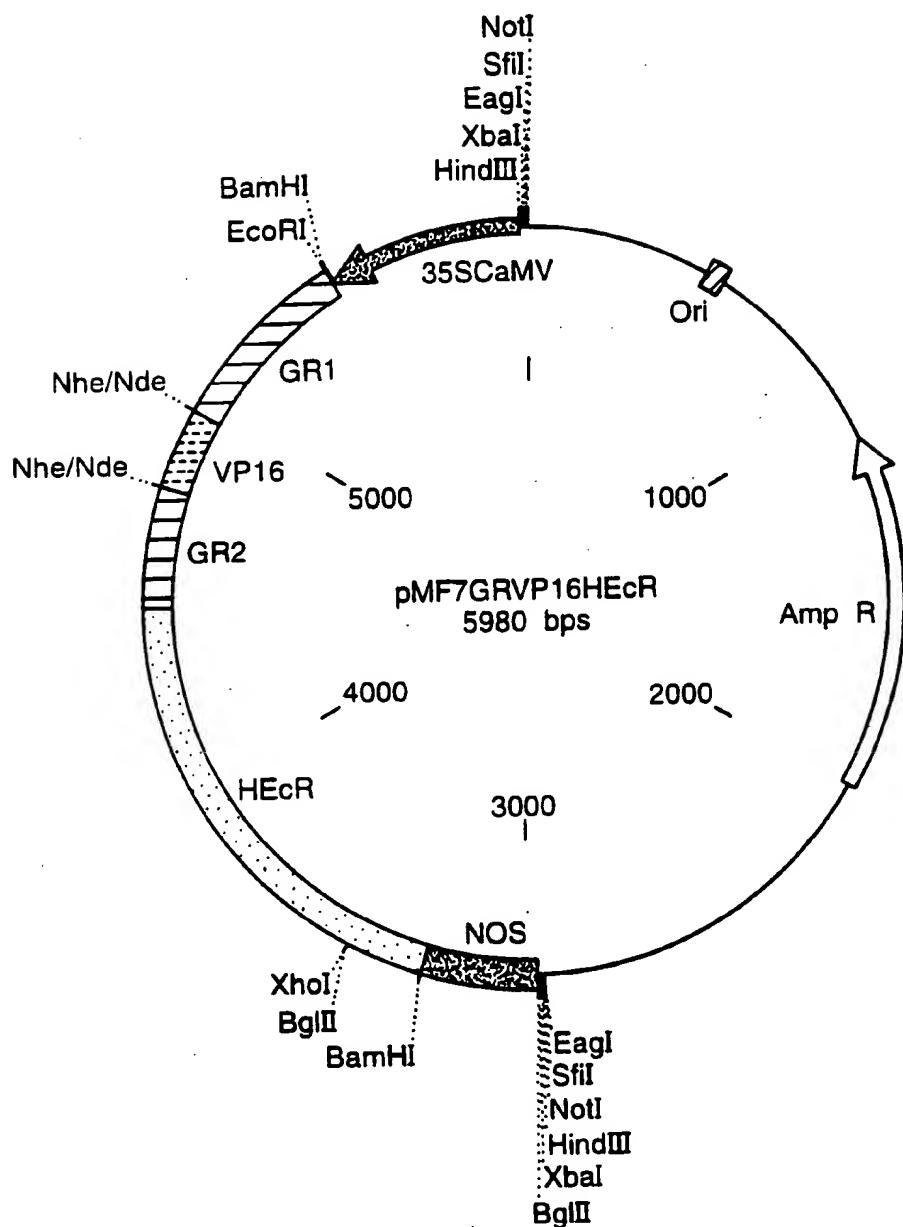
48
56

Fig.36.



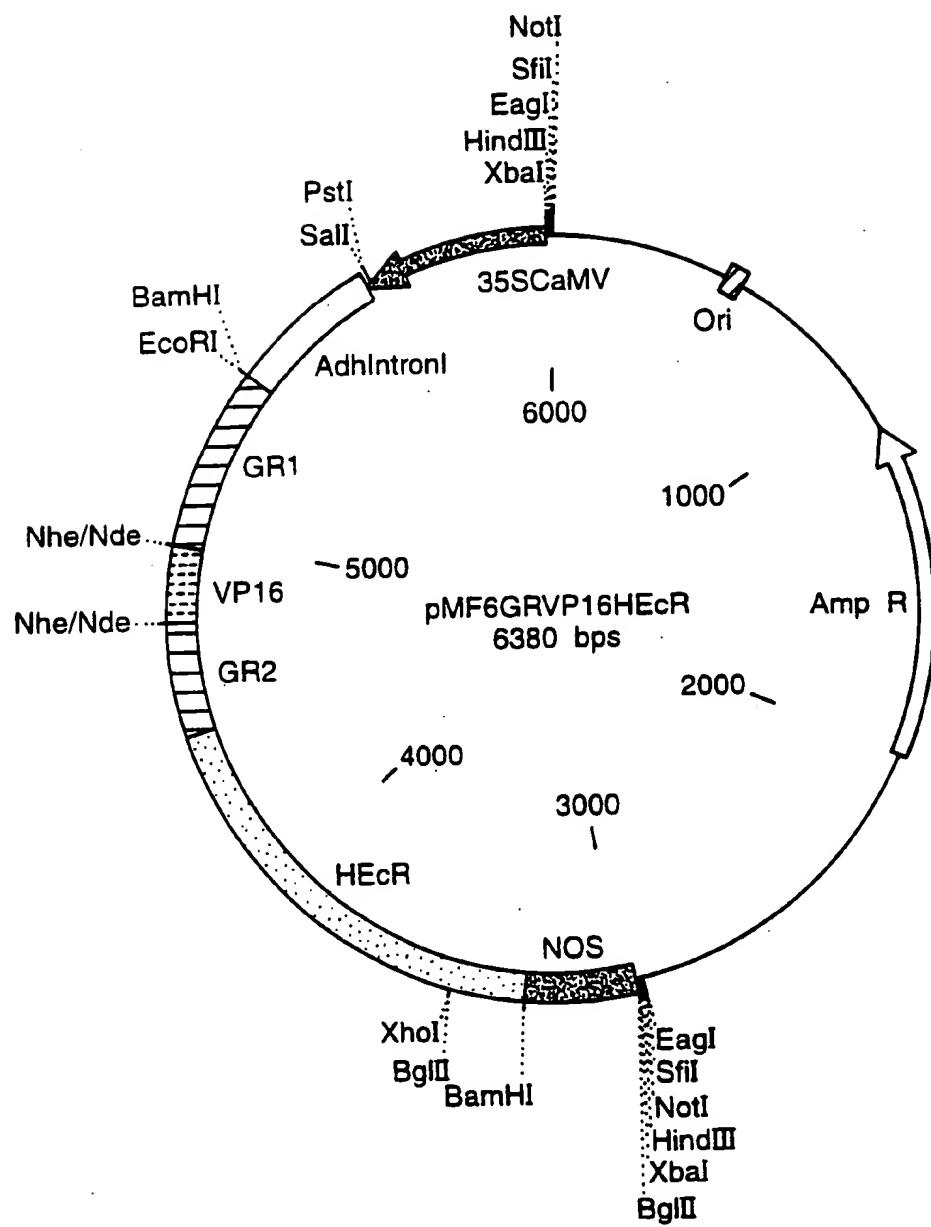
49
56

Fig.37.



50
56

Fig.38.



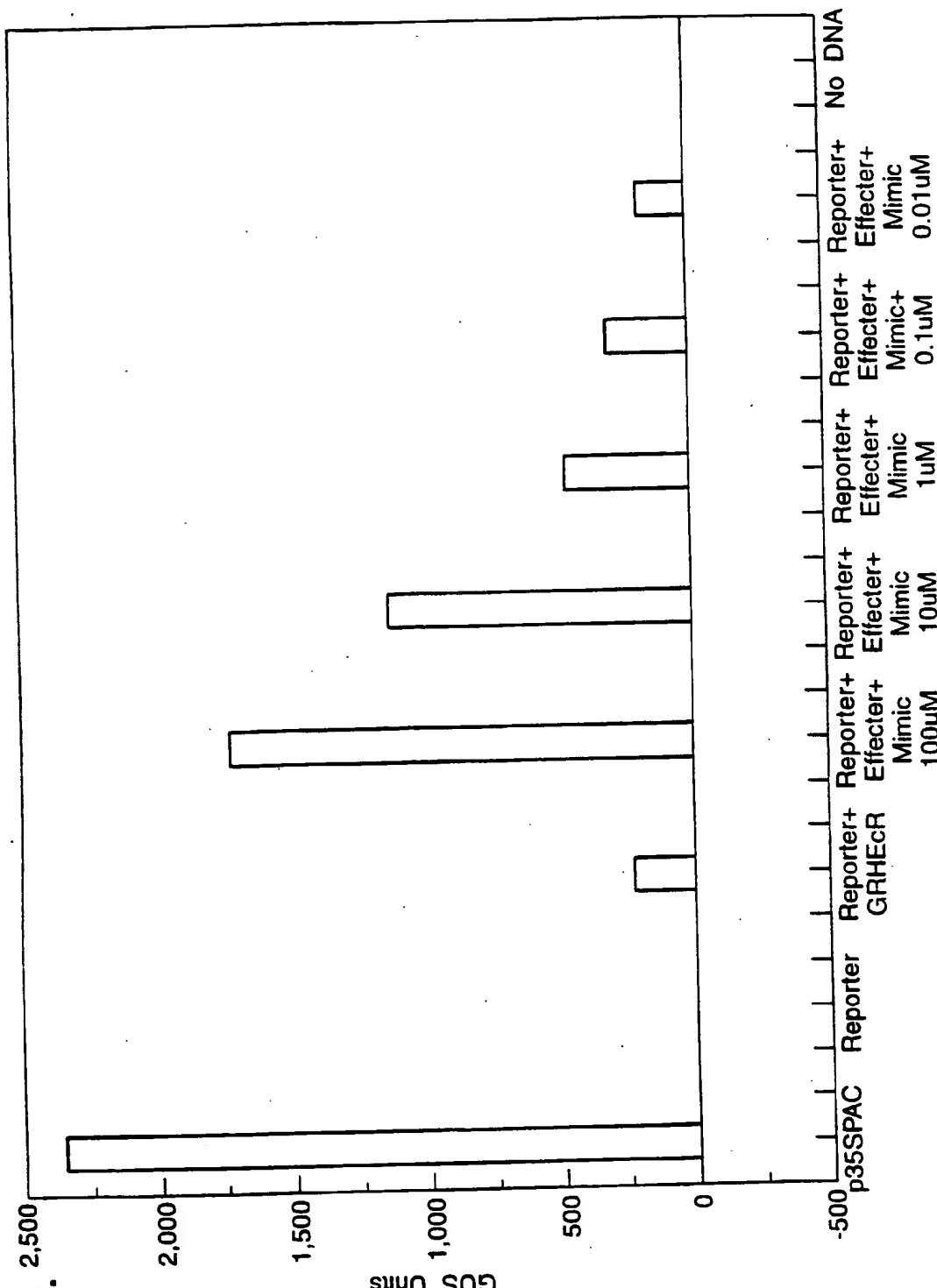
5/
56

Fig.39.

Fig.40.

Spodoptera exigua DNA sequence.
Sequence ID 6

SPODOPTERA EXIGUA HINGE AND LIGAND BINDING DOMAINS

3	9	15	21	27	33	39	45
1	AGG CCG GAG TGC GTG GTG CCA GAA AAC CAG TGT GCA ATG AAA AGG						
TCC	GGC CTC ACG CAC CAC GGT CTT GTC ACA CGT TAC TTT TCC						
46	AAG AAA AAG GCA CAA AGG GAA AAA GAC AAG TTG CCA GTC AGT						
TTT	CTC TTC CGT GTT TCC CTT CCT TTT CTG TTC AAC GGT CAG TCA						
91	ACA ACG ACA GTG GAT GAT CAC ATG CCT CCC ATT ATG CAG TGT GAT						
TGT	TGC TGT CAC CTA CTA GTG TAC GGA GGG TAA TAC GTC ACA CTA						
136	CCA CCG CCT CCA GAG GCC GCA AGA ATT CAC GAG GTG GTG CCA CGA						
GGT	GGC GGA GGT CTC CGG CGT TCT TAA GTG CTC CAC CAC GGT GCT						
181	TTC CTG AAT GAA AAG CTA ATG GAC AGG ACA AGG CTC AAG AAT GTG						
AAG	GAC TTA CTT TTC GAT TAC CTG TCC TGT TCC GAG TTC TTA CAC						
226	CCC CCT CAC TGC CAA CCA GAA GTC CTT ATT AGC GAG GCT GGT CTG						
GGG	GGA GTG ACC GTT GGT CTT CAG GAA TTA TCG CTC CGA CCA GAC						
271	GTA CCA AGA AGG CTA TGA ACA GCC ATC AGA AGA TCT AAA AAG						
CAT	GGT TCT TCC GAT ACT TGT CGG TAG TCT CCT AGA TTT TTC						

Fig.40 i.

316 AGT CAC ACA GTC GGA TGA AGA CGA AGA GTC GGA CAT GCC GTT
 TCA GTG TGT CAG CCT ACT TCT GCT TCT CCT CCT GCA CGG CAA
 361 CCG TCA GAT CAC CGA GAT GAC GAT CCT CAC AGT GCA GCT CAT TGT
 GGC AGT CTA GTG GCT CTA CTG CTA GGA GTG TCA CGT CGA GTC
 406 TGA ATT CGC TAA GGG CCT ACC AGC GTT CGC AAA GAT CTC ACA GTC
 ACT TAA GCG ATT CCC GGA TGG TCG CAA GCG TTT CTA GAG TGT CAG
 451 GGA TCA GAT CAC ATT ATT AAA GGC CTG TTC GAG TGA GGT GAT GAT
 CCT AGT CTA GTG TAA TAA TTT CCG GAC AAG CTC ACT CCA CTA CTA
 496 GTT GCG AGT AGC TCG GCG GTA CGA CGC GGC GAC AGA CAG CGT GTT
 CAA CGC TCA TCG AGC CGC CAT GCT GCG CCG CTG TCT GTC GCA CAA
 541 GTT CGC CAA CCA CCA GGC GTA CAC CCC CGA CAA CAA CCA CAA GGC
 CAA GCG GTT GTT GGT CGG CAT GTG GGC GCT GTT GAT GGC GTT CGG
 586 AGG CAT GGC CTA CGT CAT CGA GGA CCT GCT GCA CTT CTG CCG GTG
 TCC GTA CCG GAT GCA GTA GCT CCT GGA CGA CGT GAA GAC GGC CAC
 631 CAT GTA CTC CAT GAT GGA TAA CGT CCA CTA TGC ACT GCT CAC
 GTA CAT GAG GTA CTA CTC CCT ATT GCA GGT GAT ACG TGA CGA GTG
 676 TGC CAT CGT CAT TTT CTC AGA CCG ACC CGG GCT TGA GCT AAC CCT
 ACG GTA GCA GTA AAA GAG TCT GGC TGG GCC CGA ACT CGA TTG GGA
 721 GTT GGT GGA GGA GAT CCA GAG ATA TTA CCT GAA CAC GCT GCG GGT
 CAA CCA CCT CCT CTA GGT CTC TAT AAT GGA CTT GTG CGA CGC CCA

54
56

Fig.40 ii. 766 GTA CAT CCT GAA CCA GAG TCG GTC GCC GTG CTG CCC TGT CAT
CAT GTA GGA CTT GGT CTC AGC CAG CGG CAC GAC GGG ACA GTA

811 CTA CGC TAA GAT CCT CGG CAT CCT GAC GGA GCT GCG GAC CCT GGG
GAT GCG ATT CTA GGA GCC GAA GCA CCT CGT CCT CGA CGC CTG GGA CCC

856 CAT GCA GAA CTC CAA CAT GTG CAT CTC ACT CAA GCT GAA GAA CAG
GTA CGT CTT GAG GTT GTA CAC GTC GAG TGA GTT CGA CTT CTT GTC

901 GAA CGT GCC GCC GTT CTT CGA GGA TAT CTG GGA CGT CCT CGA GTA
CTT GCA CGG CGG CAA GAA CCT CCT ATA GAC CCT GCA GGA GCT CAT

946 AAA
TTT

Total number of bases is: 948.

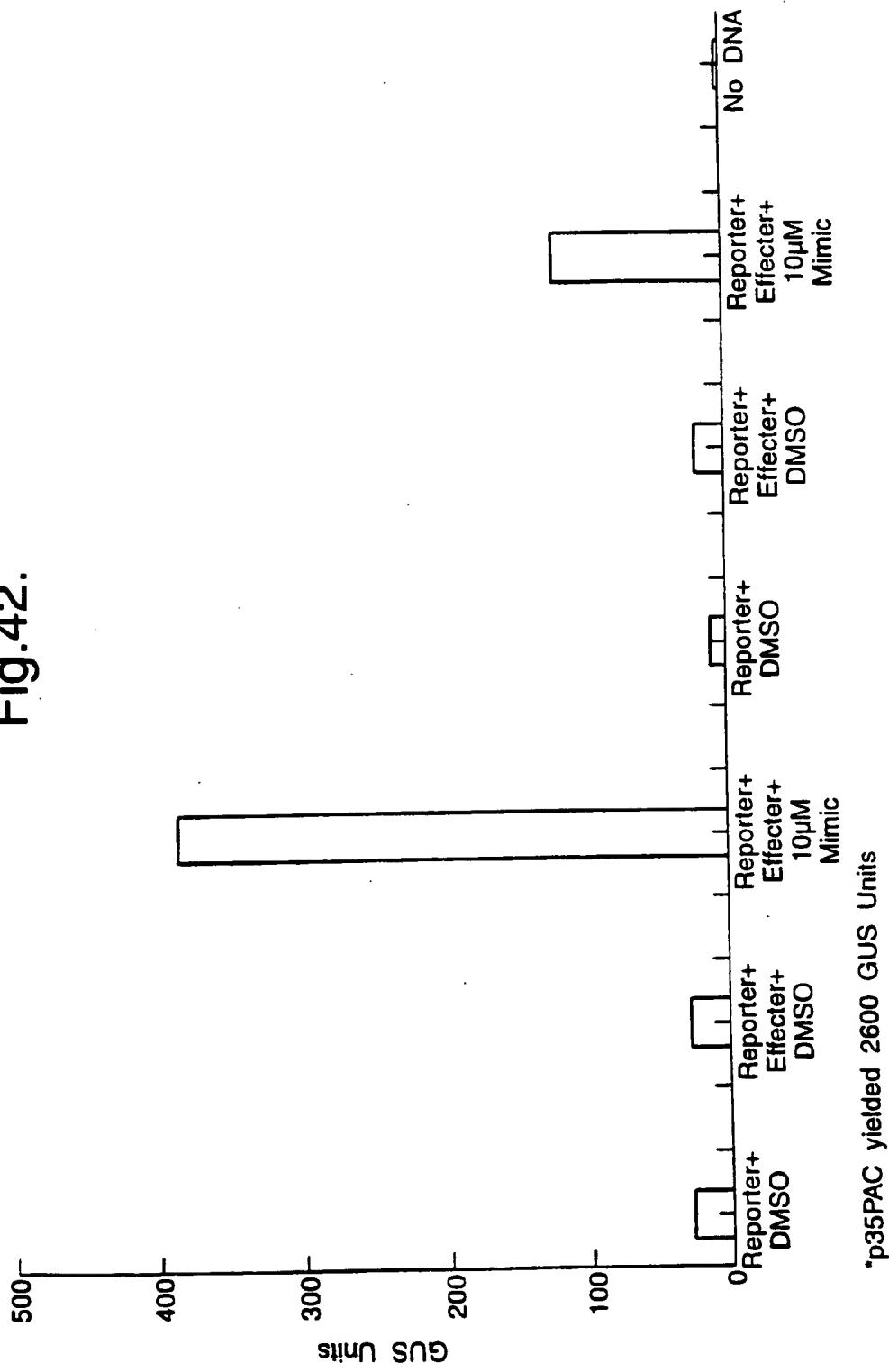
55
56

Sequence I.D. 7

Sequence comparison between *Heliocassis* 19R clone and SEC R Taq clone

HECR	RPECVVPENQCAMKRKEKKAQREKDKLKPVSTTTVDDHMPPIMQCDPPPPEAARILECVQ
SEC R	RPECVVPENQCAMKRKEKKAQREKDKLKPVSTTTVDDHMPPIMQCDPPPPEAARI
HECR	HEVVVPRFLNEKLMEQNQLKVNVPPLTANQKSLIARLVWYQEGYEQPSEEDLKRVTQSD
SEC R	HEVVVPRFLNEKLME <ins>T</ins> RLRNVPPLTANQKSLIARLVWYQEGYEQPSEEDLKRVTQSD
HECR	EDDEDSDDMPFRQITEMTILTQLIVEFAKGLP <ins>G</ins> FAKISQSDQITLLKACSSSEVMLR
SEC R	EDEEESDDMPFRQITEMTILTQLIVEFAKGLP <ins>A</ins> FAKISQSDQITLLKACSSSEVMLR
HECR	VARRYDAATDSVLFANNQAYTRDNYRKAGMAYVIEDLILHFRCRMYSMMMDNVHYALL
SEC R	VARRYDAATDSVLFANNQAYTRDNYRKAGMAYVIEDLILHFRCRMYSMMMDNVHYALL
HECR	TAIVIFSDRPGLEQPLLVEEIQRYYLNTLRVYILNQNSASPRGAVIFGEILGILTEI
SEC R	TAIVIFSDRPGLE <ins>T</ins> LLVEEIQRYYLNTLRVYILNQNSRSPCCPVIYAK <ins>K</ins> IGILTEI
HECR	RTLGMONSNMCISLKLKKRKLPPFLEIDWDV
SEC R	RTLGMONSNMCISLKLK <ins>N</ins> RNVPPFF <ins>E</ins> DIDWDV

Fig.42.



INTERNATIONAL SEARCH REPORT

Int'l Application No
PCT/GB 96/01195

A. CLASSIFICATION OF SUBJECT MATTER
 IPC 6 C12N15/12 C12N15/85 C12N15/62 C07K14/72 C07K19/00
 C12N5/10 A61K38/16

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
 IPC 6 C07K C12N A01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO,A,93 03162 (GENENTECH INC) 18 February 1993	4,5,44, 92-99
Y	see abstract; claims 1-27; figure 1	1,3, 8-43, 45-49, 51-91
X	WO,A,91 13167 (UNIV LELAND STANFORD JUNIOR) 5 September 1991	4,5,44, 50,93-99
Y	see abstract; claims 2,24 ---	2,3 -/-

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

* Special categories of cited documents :

- 'A' document defining the general state of the art which is not considered to be of particular relevance
- 'E' earlier document but published on or after the international filing date
- 'L' document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- 'O' document referring to an oral disclosure, use, exhibition or other means
- 'P' document published prior to the international filing date but later than the priority date claimed

- 'T' later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- 'X' document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- 'Y' document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

4

Date of the actual completion of the international search

9 August 1996

Date of mailing of the international search report

19.08.96

Name and mailing address of the ISA
 European Patent Office, P.B. 5818 Patentlaan 2
 NL - 2280 HV Rijswijk
 Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl.
 Fax: (+ 31-70) 340-3016

Authorized officer

Gurdjian, D

INTERNATIONAL SEARCH REPORT

Int'l Application No
PCT/GB 96/01195

C(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CELL, OCT 4 1991, 67 (1) P59-77, UNITED STATES, XP002010069 KOELLE MR ET AL: "The Drosophila EcR gene encodes an ecdysone receptor, a new member of the steroid receptor superfamily." see the whole document	4,5
Y		1-3, 8-43, 45-49, 51-92
X	INSECT BIOCHEM MOL BIOL, JAN 1993, 23 (1) P115-24, ENGLAND, XP002010070 IMHOF MO ET AL: "Cloning of a Chironomus tentans cDNA encoding a protein (cEcRH) homologous to the Drosophila melanogaster ecdysteroid receptor (dEcR)." see the whole document	4,5
X	INSECT BIOCHEM MOL BIOL, JAN 1995, 25 (1) P19-27, ENGLAND, XP002010071 CHO WL ET AL: "Mosquito ecdysteroid receptor: analysis of the cDNA and expression during vitellogenesis." see the whole document	4,5,52, 53
Y	EP,A,0 615 976 (AMERICAN CYANAMID CO) 21 September 1994 see page 6, line 28 - line 32; claims 1-12; example 2	8-43, 45-49, 51-92
Y	EUR. J. ENTOMOL. (1995), 92(1), 333-40 CODEN: EJENE2;ISSN: 1210-5759, XP002010346 SMAGGHE, GUY ET AL: "Biological activity and receptor -binding of ecdysteroids and the ecdysteroid agonists RH-5849 and RH-5992 in imaginal wing discs of Spodoptera exigua (Lepidoptera: Noctuidae)" see page 336, paragraph 3 - page 337, paragraph 2	51-65
A	DEVELOPMENTAL GENETICS, 1995, 17, 319-330, XP002010345 KOTHAPALLI R ET AL: "CLONING AND DEVELOPMENTAL EXPRESSION OF THE ECDYSONE RECEPTOR GENE FROM THE SPRUCE BUDWORM, CHORISTONEURA-FUMIFERANA" see the whole document	1-5, 51-54
		-/-

INTERNATIONAL SEARCH REPORT

Int'l Application No
PCT/GB 96/01195

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	INSECT BIOCHEM. MOL. BIOL. (1994), 24(8), 763-73 CODEN: IBMBES; ISSN: 0965-1748, XP002010072 JINDRA, MAREK ET AL: "Isolation and developmental expression of the ecdysteroid-induced GHR3 gene of the wax moth <i>Galleria mellonella</i> " see the whole document ---	1-5
A	US,A,5 424 333 (WING KEITH D) 13 June 1995 see column 150, paragraph 3 - paragraph 7; example 3 -----	97,98

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int'l Application No
PCT/GB 96/01195

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO-A-9303162	18-02-93	EP-A- 0598011 JP-T- 7501928		25-05-94 02-03-95

WO-A-9113167	05-09-91	AU-B- 1779295 AU-B- 7492291 CA-A- 2076386 EP-A- 0517805 US-A- 5514578		14-09-95 18-09-91 27-08-91 16-12-92 07-05-96

EP-A-0615976	21-09-94	CA-A- 2112445 JP-A- 6253849		01-07-94 13-09-94

US-A-5424333	13-06-95	US-A- 5354762 US-A- 4985461 AU-B- 637573 AU-B- 3636089 DE-D- 68909548 DE-T- 68909548 EP-A- 0361645 ES-T- 2059755 JP-A- 2152922 AU-B- 628349 AU-B- 3645489 DE-D- 68908789 DE-T- 68908789 EG-A- 18874 EP-A- 0347216 ES-T- 2059754 IL-A- 90606 JP-A- 2042049 PT-B- 90863 SG-A- 114793 US-A- 5117057 AU-B- 595303 AU-B- 7147387 CA-A- 1310641 DE-A- 3774347 EG-A- 18544 EP-A- 0253468 ES-T- 2037709		11-10-94 15-01-91 03-06-93 21-12-89 04-11-93 10-03-94 04-04-90 16-11-94 12-06-90 17-09-92 21-12-89 07-10-93 14-04-94 30-06-94 20-12-89 16-11-94 31-07-95 13-02-90 31-01-95 10-06-94 26-05-92 29-03-90 04-02-88 24-11-92 12-12-91 30-01-94 20-01-88 16-07-96

INTERNATIONAL SEARCH REPORT

International application No.

PCT/GB96/01195

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.: 98 because they relate to subject matter not required to be searched by this Authority, namely:
Although this claim is directed partly to a method of treatment of the human/animal body the search has been carried out and based on the alleged effects of the compound/composition
2. Claims Nos.: because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- The additional search fees were accompanied by the applicant's protest.
 No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int'l Application No
PCT/GB 96/01195

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US-A-5424333		IE-B- 59964	04-05-94
		JP-B- 7098806	25-10-95
		JP-A- 63023866	01-02-88
		KR-B- 9505199	19-05-95
		AU-B- 602505	18-10-90
		AU-B- 6926687	03-09-87
		CA-A- 1295618	11-02-92
		EP-A- 0234944	02-09-87
		ES-T- 2032818	16-07-96
		KR-B- 9410277	22-10-94
		AU-B- 599970	02-08-90
		AU-B- 7147287	31-03-88
		CA-A- 1331189	02-08-94
		DE-A- 3783111	28-01-93
		EP-A- 0261755	30-03-88
		ES-T- 2053535	01-08-94
		IE-B- 59962	04-05-94
		JP-B- 8005854	24-01-96
		JP-A- 63083063	13-04-88
		KR-B- 9513856	17-11-95
		AU-B- 597912	14-06-90
		AU-B- 6428986	30-04-87